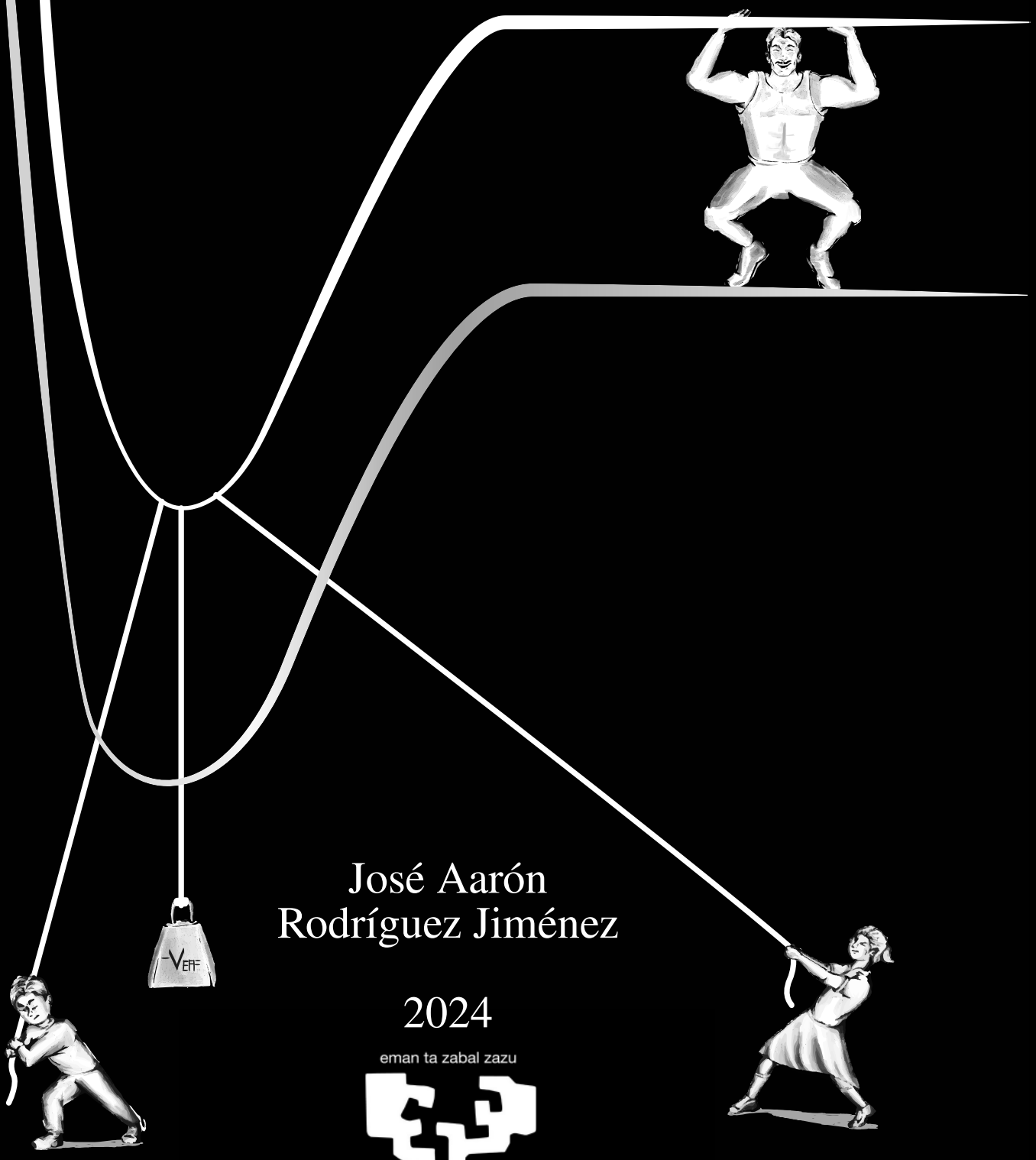
Development, implementation and applications of hybrid quantum chemistry models.



José Aarón
Rodríguez Jiménez

2024

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea





PhD Dissertation

Doctoral Program in Theoretical Chemistry and Computational Modeling

Development, implementation and applications of hybrid quantum chemistry methods.

José Aarón Rodríguez Jiménez

Advisors

Dr. David Casanova Casas

Dr. Eduard Matito

February 15, 2024

Universidad del País Vasco | Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

TESI ZUZENDARIAREN BAIMENA TESIA AURKEZTEKO	AUTORIZACIÓN DEL/LA DIRECTORA/A DE TESIS PARA SU PRESENTACIÓN

Zuzendariaren izen-abizenak /Nombre y apellidos del/la director/a: David Casanova Casas IFZ /NIF: 43631412K y Eduard Matito Gras IFZ /NIF: 40357178K.

Tesiaren izenburua / Título de la tesis: Development, implementation and applications of hybrid quantum chemistry models.

Doktorego programa / Programa de doctorado: Doctorado en Química Teórica y Modelización Computacional.

Doktoregaiaren izen-abizenak / Nombre y apellidos del/la doctorando/a: José Aarón Rodríguez Jiménez.

Unibertsitateak horretarako jartzen duen tresnak emandako ANTZEKOTASUN TXOSTENA ikusita, baimena ematen dut goian aipatzen den tesia aurkez dadin, horretarako baldintza guztiak betetzen baititu.	Visto el INFORME DE SIMILITUD obtenido de la herramienta que a tal efecto pone a disposición la universidad, autorizo la presentación de la tesis doctoral arriba indicada, dado que reúne las condiciones necesarias para su defensa.

Tokia eta data / Lugar y fecha: Donostia-San Sebástian a 01 de Enero de 2024

Sin. / Fdo.: Tesiaren zuzendaria / El/La director/a de la tesis

David Casanova Casas

Eduard Matito Gras

Universidad del País Vasco Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

AUTORIZACION DEL TUTOR/A DE TESIS
PARA SU PRESENTACION

Dr/a. José Javier López Pestaña

como Tutor/a de la Tesis Doctoral: Development, implementation and application of hybrid quantum chemistry models

realizada en el Programa de Doctorado Química Teórica y Modelización Computacional

por el Doctorando Don/ña. José Aarón Rodríguez Jiménez,

y dirigida por el Dr./a David Casanova Casas y por el Dr./a Eduard Matito Gras.

autorizo la presentación de la citada Tesis Doctoral, dado que reúne las condiciones necesarias para su defensa.

En Donostia-San Sebástian a 11 de Diciembre de 2023

EL/LA TUTOR/A DE LA TESIS

Fdo.: Firmado por LOPEZ PESTAÑA JOSE JAVIER - ***9823** el día

Universidad del País Vasco Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

AUTORIZACIÓN DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO

La Comisión Académica del Programa de Doctorado en Química Teórica y Modelización Computacional en reunión celebrada el día 12 de Diciembre de 2023, ha acordado dar la conformidad a la presentación de la Tesis Doctoral titulada:

Development, implementation and applications of hybrid quantum chemistry models

dirigida por el Dr. David Casanova Casas y por el Dr. Eduard Matito Gras. y presentada por Don José Aarón Rodríguez Jiménez adscrito o adscrita al Departamento de Polímeros y Materiales Avanzados (PMAS)

En Donostia-San Sebástian a 12 de Diciembre de 2023

EL/LA RESPONSABLE DEL PROGRAMA DE DOCTORADO

Firmado por LOPEZ PESTAÑA JOSE JAVIER - ***9823** el día 13/12/2023 con un certificado emitido por

Fdo.: _____

Universidad del País Vasco | Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

AUTORIZACIÓN DEL DEPARTAMENTO

El Consejo del Departamento de Polímeros y Materiales Avanzados (PMAS)

en reunión celebrada el día _16__ de _enero___ de __2024__ ha acordado dar la conformidad a la

admisión a trámite de presentación de la Tesis Doctoral titulada:

Development, implementation and applications of hybrid quantum chemistry models

dirigida por el/la Dr/a. David Casanova Casas

y por el/la Dr/a. Eduard Matito Gras

y presentada por Don/ña. José Aarón Rodríguez Jiménez

ante este Departamento.

En __Donostia-San Sebastián____ a ___16__ de _____enero_____ de __2024___

VºBº DIRECTOR/A DEL DEPARTAMENTO

AGUSTIN ETXEBERRIA LIZARRAGA - 72434443Z

Firmado digitalmente por AGUSTIN ETXEBERRIA LIZARRAGA - 72434443Z
Fecha: 2024.01.16 17:22:52 +01'00'

Fdo.: ___Agustin Exteberria Lizarraga__

SECRETARIO/A DEL DEPARTAMENTO

JOSEBA IÑAKI JUARISTI OLIDEN - 34089368X

Firmado digitalmente por JOSEBA IÑAKI JUARISTI OLIDEN - 34089368X
Nombre de reconocimiento (DN): c=ES, o=UPV - EHU, ou=Ziurtagiri Profesionala - Certificado Profesional, 2.5.4.97=VATES-Q4818001B, cn=JOSEBA IÑAKI JUARISTI OLIDEN - 34089368X, givenName=JOSEBA IÑAKI, sn=JUARISTI OLIDEN, serialNumber=34089368X
Fecha: 2024.01.16 13:34:58 +01'00'

Fdo.: ____Joseba Iñaki Juaristi Oliden

Universidad del País Vasco | Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

ACTA DE GRADO DE DOCTOR O DOCTORA
ACTA DE DEFENSA DE TESIS DOCTORAL

DOCTORANDO/A DON/DÑA. José Aarón Rodríguez Jiménez

TITULO DE LA TESIS: Development, implementation and applications of hybrid quantum chemistry models.

El Tribunal designado por la Comisión de Postgrado de la UPV/EHU para calificar la Tesis Doctoral arriba indicada y reunido en el día de la fecha, una vez efectuada la defensa por el/la doctorando/a y contestadas las objeciones y/o sugerencias que se le han formulado, ha otorgado por_____la calificación de:

unanimidad ó mayoría

SOBRESALIENTE / NOTABLE / APROBADO / NO APTO

Idioma/s de defensa (en caso de más de un idioma, especificar porcentaje defendido en cada idioma):

Castellano _____

Euskera _____

Otros Idiomas (especificar cuál/cuales y porcentaje) _____

En _____a _____de _____de _____

EL/LA PRESIDENTE/A, EL/LA SECRETARIO/A,

Fdo.: Fdo.:

Dr/a: _____ Dr/a: _____

VOCAL 1º, VOCAL 2º, VOCAL 3º,

Fdo.: Fdo.: Fdo.:

Dr/a: _____Dr/a:_____Dr/a:_____

EL/LA DOCTORANDO/A,

Fdo.: _____

Esker onak / Acknowledgements

Agradezco a las facilidades de cómputo, al apoyo técnico y humano recibido por el Donostia International Physics Center (DIPC).

Al Gobierno Vasco por la financiación económica (proyecto PIBA19-0004).

Al apoyo de Movilidad del DIPC que contribuyó a que pudiese llevar a cabo una estancia corta de investigación en el grupo de la Dra. Shirin Faraji en el instituto Zernike para materiales avanzados de la Universidad de Groningen (Groningen, Paises Bajos).

Agradezco enormemente a mis asesores, David Casanova y Eduard Matito, quienes se han convertido en personas de gran estima personal y me han guiado con una paciencia extraordinaria por la larga senda de la Ciencia, me han apoyado y siempre han visto por mi.

En su día le pregunté a Juvencio Robles, "Bueno, hacemos mucho cálculo, obtenemos las energías, las geometrías ¿y luego qué? ¿para qué nos sirve?" Nunca imaginé que su respuesta me llevaría a emprender esta gran aventura. Antes de comenzar el master le dije a Faustino Aguilera "¿Y yo? ¿Qué voy a estar haciendo ahí?, son físicos y yo químico, apenas sé algo de física." El me respondió, "si, pero tu tienes habilidades que ellos tampoco tienen". Juve, Faust, estoy enormemente agradecido por todo su apoyo y por aparecer en mi vida. Gracias por creer en mí.

A mí madre, Liz, por su infinita paciencia y amor, que me ha inspirado desde siempre, y que de alguna forma hace que vuelva a poner los pies en la tierra. Gracias por las fuerzas que me diste para seguir adelante. A mi padre, Juan, ojalá pudieras leer esto. A mi hermana, Abi, después de todo gracias a tí decidí dedicarme a esta area del conocimiento.

Después de escribir dos tesis y revisando lo que escribí en su momento, he notado mi crecimiento personal. El doctorado, para mí, ha sido una maravilla, un sueño cumplido. Y en cierto sentido, esta oportunidad me ha salvado la vida. Ya desde hace varios años que lucho a diario contra una aflicción que me desgarra por dentro. Por esto mismo, no puedo evitar que mis agradecimientos adquieran matices pasionales, quisiera agradecer de corazón a quienes siendo heroes anónimos, amigos, enemigos, personas y familia han contribuido a este crecimiento. Palabras que ilustran mejor lo que quiero decir y que tomo prestadas de Fernando Savater en su Invitación a la Ética son:

'A los que me conocieron y me olvidaron, pues me dieron el trato que debemos esperar de los dioses propicios. A quienes he defraudado, a quienes han sufrido por mi causa. A mis enemigos, con agradecimiento por la atención que me prestan. A los que se han reido conmigo y de mi. A todo el que un día me dijo sinceramente: "estoy a tu lado". A las que se fueron y a la que vendrá. A mis amigos y a ti, por si quieres llegar a serlo.'

Y de mi propia cosecha pero inspiradas en vivencias previas, dedicadas a todos ustedes:

"*Las espadas más peligrosas son las que se forjan con los golpes más duros, y los resisten.*
Los diamantes más hermosos solo pueden formarse bajo presión.
Pero no olvidar que, del estiércol nacen flores.
Cosa que ni la letalidad de la espada, ni la belleza del diamante pueden ofrecer."

Para ustedes, ya sabéis quienes sois.

Abstract

The field of theoretical quantum chemistry aims to organize chemical laws, principles and rules by applying quantum mechanics to various challenges within the real of chemistry. Central to this framework is the interconnection of the structure with the properties of molecular systems. This interconnection is used to provide theories and explanations for chemical phenomena, such as bonding, heats of formation, and reactivity mechanisms, among others.

From the early days of quantum mechanics, electronic structure theory establishes the fundamental laws necessary for the mathematical treatment of chemistry problems. These laws are well-established and understood. However, the enormous task of method development is oriented towards overcoming the limitations imposed by the large computational cost involved in the solutions to the many-body electronic Schrödinger equation.

Achieving accurate or exact solutions for this equation allows for a complete description of the motions of electrons in atoms and molecules, posing a formidable challenge. Although theoretically feasible, in practice, the exact solution is prohibitively expensive to compute for large systems. Thus, designing a method that delivers cheap and accurate approximate solutions is primordial for the development of theoretical chemistry and to extend accurate solutions to large systems. In developing such methods, the description of electronic correlation is of fundamental importance. This refers to the influence that the motion of one electron has on the motion of another electron in a many-electron system. Central to this description is the concept of electron correlation energy, defined as a measure for the energy difference between the exact solution and that captured by mean field models (e.g., Hartree-Fock).

Electron correlation can be classified into the so-called dynamic and nondynamic correlation types. While the former is considered universal in the sense that it is associated with the instantaneous electron-electron interaction, the latter is related with system's attributes, such as the relative energy of the orbitals (the mathematical functions used to represent the electrons in a molecule). In fact, most of available theoretical treatments can accurately handle just one or the other of these two correlation types. While there are options for simultaneously addressing both types of correlation, their widespread adoption is impeded by the substantial computational workload associated with these approaches.

Scientific research works included in this thesis are focused to merge different strategies, suitable for simultaneously treat dynamic and nondynmic correlation, together in an unified framework. Such that a balanced treatment of electronic correlation can be achieved with an affordable computational cost. This thesis addresses the development and implementation of theoretical chemistry methodologies designed to account for a balanced treatment of electron correlation within an accessible computational cost, and investigates their performance and limitations.

Initially, we explore the capabilities of wave function theory (WFT) methodologies with density functional theory (DFT), resulting in improved mingled WFT-DFT approximations. We select the WFT part to handle the treatment of non-dynamic correlation, while the remaining dynamic correlation is incorporated from DFT. One of the common strategies employed to merge WFT with DFT methods relies on the range-separation of the Coulomb operator, in which DFT functionals take care of the short-distance part, while long-range inter-electronic interactions are evaluated by the chosen wave function method (WFT-*sr*DFT). We uncover the limitations of WFT-*sr*DFT in the characterization of open-shell systems. We show that spin polarization effects have a major impact on the (short-range) DFT exchange energy and are of vital importance to provide a balanced description between closed and open-shell configurations. Furthermore, we introduce and test different strategies to account for spin polarization in the short-range. Our results indicate that the use of short-range DFT correlation (without exchange) in combination with a (full-range) multiconfigurational wave function might be an excellent approach for the study of open-shell molecules, largely improving the performance of WFT and WFT-*sr*DFT.

Moreover, we introduce a novel hybrid approach, termed WFT-*so*DFT, in which the coupling between the two theories is performed through a partitioning of the orbital space, distinguishing between large and small natural occupation numbers (NOONs) associated with WFT and DFT correlation, respectively. To this end, we employ a local density correlation functional aiming to only account for small-occupation correlation energy. Our results exemplify the method's ability to handle small but challenging chemical systems, for which WFT-*so*DFT shows a notable improvement over the pristine WFT method. However, we identify problems associated with the continuity of the occupations along dissociation curves, the size of the basis set and the choice of the parameter associated with the truncation of the orbital space. Our findings collectively highlight the potential of the WFT-*so*DFT approach as a computationally affordable strategy to handle electron correlations.

We further explore an alternative method, consisting in the use of a two-electron effective potential able to retrieve the dynamic correlation contribution to a complete active space self consistent field (CASSCF) wave function. The effective potential depends on the electronic separation, and resembles the form of a Gaussian-damped Coulomb interaction. Through the study of atomic and diatomic systems, we demonstrate the limitations and strengths of the proposed approach. FCI quality results can be achieved within a proper parametrization of the effective potential. This turns to be one of the major challenges in the development of a general approach of this type. In order to rationalize the role played by the parameters in the potential, we analyse changes in energy, electron correlation and electron distribution. To this end, Coulomb hole analysis and electron correlation indexes show their relevance in the development of the method.

From Coulomb hole analysis made for representative systems, we devise the main effects delivered by the effective potential into the CASSCF wave function. Drawing from our findings, we propose a modified approach involving the segmentation of the effective potential into inte- and intra-atomic components. We then investigate distinct impacts of these contributions. Accurate results obtained with this approach in the calculation of diatomic potential energy surfaces of molecules with different spin states points to promising insights for future works.

Finally, we turn our attention to the applicability of a mixed scheme that incorporates spin-orbit coupling contributions in a perturbative way. Such calculations are based on the application of the spin-orbit mean-field approximation with non-relativistic wave functions.

Within this framework, we perform electronic structure calculations to unravel the origin of SOCs in model open-shell molecules, with special focus on difficult scenarios such as molecular distortions and intermolecular interactions. Obtained results demonstrate the suitability of our WFT method in the calculation of SOCs of open-shell systems, while providing a deep understanding of the relationship between couplings and the collective nature of the electronic states. During this study, we introduce a new definition of the SOC constant for the study of molecular aggregates.

Resumen

El campo de la química cuántica teórica tiene como objetivo organizar leyes, principios y reglas químicas aplicando la mecánica cuántica a diversos desafíos dentro del ámbito de la química. Fundamental en este marco es la interconexión entre la estructura y las propiedades de los sistemas moleculares. Esta interconexión se utiliza para proporcionar teorías y explicaciones de fenómenos químicos, como la formación de enlaces, los calores de formación y los mecanismos de reactividad, entre otros.

Desde los primeros días de la mecánica cuántica, las leyes fundamentales necesarias para el tratamiento matemático de problemas químicos se han consolidado en el marco de la teoría de la estructura electrónica. Estas leyes están bien establecidas y comprendidas. Sin embargo, la enorme tarea del desarrollo de métodos que queda, se ha orientado a superar las limitaciones impuestas por el alto costo computacional involucrado en las soluciones de la ecuación de Schrödinger electrónica de muchos cuerpos.

Lograr soluciones exactas o precisas para esta ecuación permite una descripción completa de los movimientos de los electrones en átomos y moléculas, presentando un desafío formidable. Aunque teóricamente factible, en la práctica, la solución exacta es prohibitivamente costosa de calcular para sistemas con varios átomos. Por lo tanto, diseñar un método que proporcione soluciones aproximadas económicas y precisas es primordial para el desarrollo de la química teórica y para extender soluciones precisas a sistemas grandes. En el desarrollo de tales métodos, la descripción de la correlación electrónica es de importancia fundamental. Esto se refiere a la influencia que tiene el movimiento de un electrón en el movimiento de otro en un sistema de muchos electrones. Fundamental en esta descripción es el concepto de energía de correlación electrónica, definida como una medida de la diferencia de energía entre la solución exacta y la capturada por modelos de campo medio (por ejemplo, Hartree-Fock).

La correlación electrónica se puede clasificar en los llamados tipos de correlación dinámica y no dinámica. Mientras que la primera se considera universal en el sentido de que está asociada con la interacción instantánea electrón-electrón, la segunda está relacionada con atributos del sistema, como la energía relativa de los orbitales (las funciones matemáticas utilizadas para representar a los electrones en una molécula). De hecho, la mayoría de los tratamientos teóricos disponibles pueden manejar con precisión solo uno u otro de estos dos tipos de correlación. Aunque existen opciones para abordar simultáneamente ambos tipos de correlación, su adopción generalizada se ve obstaculizada por la carga computacional sustancial asociada con estos enfoques.

Los trabajos de investigación científica incluidos en esta tesis se centran en fusionar diferentes estrategias, adecuadas para tratar simultáneamente la correlación dinámica y no dinámica, en un marco unificado. De esta manera, se puede lograr un tratamiento equilibrado de la correlación electrónica con un costo computacional asequible. Esta tesis aborda el desarrollo e implementación de metodologías de química teórica diseñadas para tener en cuenta un

tratamiento equilibrado de la correlación electrónica dentro de un costo computacional accesible, e investiga su rendimiento y limitaciones.

Inicialmente, exploramos las capacidades de las metodologías de teoría de la función de onda (WFT) con la teoría del funcional de la densidad (DFT), lo que resulta en aproximaciones mejoradas de WFT-DFT mezcladas. Seleccionamos la parte de WFT para manejar el tratamiento de la correlación no dinámica, mientras que la correlación dinámica restante se incorpora desde DFT. Una de las estrategias comunes empleadas para fusionar los métodos WFT con DFT se basa en la separación de rango del operador de Coulomb, en la cual los funcionales de DFT se encargan de la parte de corta distancia, mientras que las interacciones electrónicas de larga distancia se evalúan mediante el método de función de onda elegido (WFT-*sr*DFT). Revelamos las limitaciones de WFT-*sr*DFT en la caracterización de sistemas de capa abierta. Mostramos que los efectos de polarización de espín tienen un impacto importante en la energía de intercambio de DFT (de corto alcance) y son de vital importancia para proporcionar una descripción equilibrada entre configuraciones de capa cerrada y abierta. Además, introducimos y probamos diferentes estrategias para tener en cuenta la polarización de espín en el corto alcance. Nuestros resultados indican que el uso de la correlación de DFT de corto alcance (sin intercambio) en combinación con una función de onda multiconfiguracional (de alcance completo) podría ser un enfoque excelente para el estudio de moléculas de capa abierta, mejorando en gran medida el rendimiento de WFT y WFT-*sr*DFT.

Además, presentamos un enfoque híbrido novedoso, denominado WFT-*so*DFT, en el cual el acoplamiento entre las dos teorías se realiza a través de una partición del espacio orbital, distinguiendo entre números de ocupación natural (NOONs) grandes y pequeños asociados con la correlación de WFT y DFT, respectivamente. Para ello, empleamos una función de correlación de densidad local con el objetivo de tener en cuenta solo la energía de correlación de ocupación pequeña. Nuestros resultados ejemplifican la capacidad del método para manejar sistemas químicos pequeños pero desafiantes, para los cuales WFT-*so*DFT muestra una notable mejora sobre el método de WFT prístino. Sin embargo, identificamos problemas asociados con la continuidad de las ocupaciones a lo largo de curvas de disociación, el tamaño del conjunto de base y la elección del parámetro asociado con la truncación del espacio orbital. Nuestros hallazgos resaltan colectivamente el potencial del enfoque WFT-*so*DFT como una estrategia computacionalmente asequible para manejar correlaciones electrónicas.

Exploramos además un método alternativo, que consiste en el uso de un potencial efectivo de dos electrones capaz de recuperar la contribución de correlación dinámica a una función de onda autoconsistente de espacio activo completo (CASSCF). El potencial efectivo depende de la separación electrónica y se asemeja a la forma de una interacción Coulombiana amortiguada por una función gaussiana. A través del estudio de sistemas atómicos y diatómicos, demostramos las limitaciones y fortalezas del enfoque propuesto. Se pueden lograr resultados de calidad FCI mediante una parametrización adecuada del potencial efectivo. Esto resulta ser uno de los principales desafíos en el desarrollo de un enfoque general de este tipo. Para racionalizar el papel desempeñado por los parámetros en el potencial, analizamos cambios en la energía, correlación electrónica y distribución electrónica. Con este fin, el análisis del agujero de Coulomb y los índices de correlación electrónica muestran su relevancia en el desarrollo del método.

A partir del análisis del agujero de Coulomb realizado para sistemas representativos, ideamos los principales efectos entregados por el potencial efectivo en la función de onda CASSCF. Basándonos en nuestros hallazgos, proponemos un enfoque modificado que implica la

segmentación del potencial efectivo en componentes inter y intraatómicos. Luego, investigamos impactos distintos de estas contribuciones. Los resultados precisos obtenidos con este enfoque en el cálculo de superficies de energía potencial diatómicas de moléculas con diferentes estados de espín apuntan a perspectivas prometedoras para trabajos futuros.

Finalmente, dirigimos nuestra atención a la aplicabilidad de un esquema mixto que incorpora contribuciones de acoplamiento espín-órbita de manera perturbativa. Dichos cálculos se basan en la aplicación de la aproximación de campo medio espín-órbita con funciones de onda no relativistas. Dentro de este marco, realizamos cálculos de estructura electrónica para desentrañar el origen de los acoplamientos espín-órbita en moléculas modelo de capa abierta, con especial atención a escenarios difíciles como distorsiones moleculares e interacciones intermoleculares. Los resultados obtenidos demuestran la idoneidad de nuestro método WFT en el cálculo de acoplamientos espín-órbita en sistemas de capa abierta, al tiempo que proporcionan una comprensión profunda de la relación entre los acoplamientos y la naturaleza colectiva de los estados electrónicos. Durante este estudio, introducimos una nueva definición de la constante de acoplamiento espín-órbita para el estudio de agregados moleculares.

Contents

Contents **xi**

List of Figures **xv**

List of Tables **xxiii**

1 Introduction **1**
1.1 Overview of this thesis . . . 3
1.2 Electron Correlation . . . 4
1.2.1 Methods for the treatment of electron correlation. . . . 6
1.2.2 Desirable properties in an electronic structure method. . . . 9

2 A review of theoretical background **13**
2.1 Hartree-Fock approximation . . . 13
2.2 Configuration interaction formalism . . . 17
2.2.1 CIPSI algorithm . . . 20
2.3 Complete active space approaches . . . 21
2.3.1 Complete active space configuration interaction (CASCI) . . . 22
2.3.2 Restricted active space configuration interaction (RASCI) . . . 25
2.3.3 Multiconfigurational Self-Consistent Field (MCSCF) . . . 27
2.4 Dynamic correlation in CI approaches . . . 28
2.5 Density functional theory . . . 29
2.5.1 Kohn-Sham Density Functional Theory . . . 30
2.6 Overview of selected methods for simultaneous treating of DC and NDC . 32
2.6.1 CASPT2 and CAS-NEVPT2 . . . 32
2.6.2 CAS-DFT . . . 35
2.6.3 DFT/MRCI . . . 37
2.6.4 MCPDFT and CASIIDFT methods . . . 38
2.6.5 Range separation formalism . . . 42
2.6.6 Adiabatic connection based methods . . . 44
2.7 Tools for analysis . . . 48
2.7.1 Natural orbitals . . . 48
2.7.2 Correlation indices based on natural orbitals . . . 50
2.7.3 Fractional occupation density . . . 51
2.7.4 Natural transition orbitals . . . 52
2.7.5 Coulson's electron correlation hole . . . 54

3 Objectives **57**

4 Short-range DFT energy correction to multiconfigurational wavefunctions for open-shell systems 59
4.1 Abstract . . . 59
4.2 Introduction . . . 59
4.3 Theory . . . 61
4.3.1 Range separation in WFT-*sr*DFT . . . 61
4.3.2 State average approximation . . . 62
4.3.3 From paired to unpaired electrons . . . 63
4.3.4 Spin polarized short-range DFT . . . 65
4.3.5 Exact short-range exchange . . . 66
4.4 Computational Details . . . 67
4.5 Numerical results . . . 67
4.5.1 Hydrogen dissociation . . . 68
4.5.2 Triplet state atoms: C, O, Si and S . . . 69
4.5.3 Triatomic diradicals: CH_2, NH_2^+, SiH_2 and PH_2^+ . . . 71
4.5.4 Benzyne isomers . . . 74
4.5.5 Trimethylenemethane diradical . . . 76
4.5.6 Gallium dimer . . . 79
4.6 Conclusions . . . 82

5 Small-occupation density functional correlation energy correction to wave function approximations 83
5.1 Abstract . . . 83
5.2 Introduction . . . 83
5.3 General framework to merge WFT with DFT . . . 85
5.4 Orbital separation based on electron occupation . . . 86
5.4.1 Occupation separation parameter . . . 87
5.4.2 Small-occupation correlation energy functionals . . . 89
5.4.3 Sampling the orbital occupation space . . . 89
5.5 Methods . . . 91
5.5.1 Computational Workflow . . . 91
5.5.2 Computational details and nomenclature . . . 92
5.6 Results and Discussion . . . 93
5.6.1 Helium and beryllium atomic series . . . 93
5.6.2 H_2 dissociation . . . 95
5.6.3 Ethylene torsion . . . 96
5.6.4 Multiple bond breaking: N_2 dissociation . . . 97
5.7 Conclusions . . . 99

6 Introducing dynamical correlation into CASSCF wave functions through an effective potential 101
6.1 Abstract . . . 101
6.2 Introduction . . . 101
6.3 Methods Based on the Correlation Hole . . . 103
6.3.1 The Roos CASSCF Effective Potential Method . . . 105
6.3.2 Drawbacks and Strengths of the Method . . . 106
6.4 Methods . . . 107
6.4.1 Computational Workflow . . . 107
6.4.2 Computational Details . . . 108

6.5 Numerical Results . . . 108
6.5.1 Performance in atomic systems . . . 108
6.5.2 Describing H_2 dissociation . . . 112
6.6 Splitting the Roos' Effective Potential . . . 115
6.6.1 Studies in diatomics beyond H_2 . . . 117
6.7 Conclusions . . . 129

7 Spin-Orbit Couplings of Open-Shell systems with Restricted Active Space Configuration Interaction 133
7.1 Abstract . . . 133
7.2 Introduction . . . 133
7.3 Theory . . . 134
7.3.1 The Breit-Pauli Hamiltonian . . . 134
7.4 Computational details . . . 136
7.5 Results and discussion . . . 136
7.5.1 Trimethylene biradical . . . 137
7.5.2 Single bond dissociation: ethylene twist . . . 140
7.5.3 Double bond dissociation: ethylene stretch . . . 143
7.5.4 SOC in the ethylene-O_2 system . . . 150
7.6 Conclusions . . . 155

8 General Conclusions 157

Appendix A 161
Additional information for H_2 . . . 161
Additional information for triatomic diradicals . . . 162

Appendix B 163
Additional information for polyene systems . . . 163
Numerical results for atomic series . . . 165
Complementary data for Atomic series . . . 165
Performance of different variants of the method . . . 166
Comparison of correlation energies computed for the Be-series. . . . 167
Total energies computed with the cc-pVQZ basis set . . . 168
H_2 bond dissociation . . . 169
Complementary data for H_2 dissociation . . . 169
Basis set dependence . . . 170
Potential energy surfaces for H_2 . . . 170
Energy dependence on ν parameter . . . 173
C_2H_4 molecular torsion . . . 174
Complementary data for C_2H_4 torsion . . . 174
Potential energy surfaces for C_2H_4 . . . 174
Energy dependence on ν parameter . . . 177
N_2 triple bond dissociation . . . 178
RASCI Natural orbital occupations computed for the N_2 . . . 178
Potential energy surfaces for N_2 . . . 179
Energy dependence on ν parameter . . . 187
Reducing discontinuities . . . 188
Discontinuities at larger basis sets . . . 189

Size consistency analysis 190

Appendix C **193**
Solution for Attenuated Electron Repulsion Integrals Damped by a Gaussian Two-electron Operator 193
Solution for the Gaussian two-electron repulsion integrals in the Boys picture. 193
Treatment of the Cartesian angular momentum components 195
Use of the Gauss product theorem 196
Solution of the inner integral (G_{RI}) I: Treatment of the attenuated Coulomb operator 197
Solution of the inner integral (G_{RI}) II: Solutions for the Cartesian components. 198
Further expansions in G_{RI} and k-space integration 202
The use of Boys function 204
Complete form of the gauERIs 206
Fundamental gauERI 207
Additional information for atomic systems 208
Correlation energies computed for atoms with different μ ν values 208
Total CASSCF+V_{eff} and corresponding correlation energies computed for cations with different μ ν values 210
Total CASSCF+V_{eff} and corresponding correlation energies computed for anions with different μ ν values 212
Coulomb holes for C atom 213
Additional information for diatomic systems 215
CASSCF+V_{eff} PES and Coulomb holes for H_2 215
Results for Isoelectronic Diatomics: N_2 and NO^+ 217
Results for Triplet State Diatomics: O_2 and NH 217
Best estimates for O_2 dissociation. 218
Best estimates for NH dissociation. 219
Results for Doublet State Diatomics: OH 219
Additional information regarding parametrization strategy 2 (PS2) and data fitting. 220

Appendix D **223**
Mean-Field Spin-Orbit Coupling 223
Comparative Results for Trimethylene 225
Complementary data for the C_2H_4 model 228
SOC in the ethylene-O_2 system 234

List of Publications **237**

Bibliography **239**

List of Figures

1.1 Some of the available DFT+MC-WFT mingled approaches. A very brief description of their features is included inside dashed-line squares. 9
1.2 Pople's five stages of the development process for a theoretical model, adapted from Ref. [1]. 10

2.1 Representation of the three RAS spaces and the action of the $\hat{r_0}$, $\hat{r_h}$ (*hole*) and $\hat{r_p}$ (*particle*) terms of the excitation operator on the reference state ($|\Psi_0\rangle$), here exemplifying a triplet state with an active space of six electrons in four orbitals (RAS2(6e,4o)). 26
2.2 (a) Ilustrative comparison of CASSCF and FCI Coulomb holes computed for the H_2 at equilibrium distance (0.77 Å), the lack of dynamic correlation in the CASSCF wave function produces a qualitative change in the shape of the hole with respect to the FCI solution. (b) Same comparison computed at large separation (3.85 Å), in this case both solutions retrieve the same hole. Calculations performed with the cc-pVDZ basis set. 56

4.1 Ground state dissociation energy curves of H_2 computed at the FCI, RAS-SF and RAS-*sr*LDA levels with the 6-31G basis set. 64
4.2 RAS-*sr*LDA dissociation energy curves of H_2 for the ground state singlet (S_0) and the T_1 with $M_S = 0$ and $M_S = \pm 1$ triplet microstates (a) and the *sr*-DFT exchange and correlation contributions (b) computed with the 6-31G basis set. 65
4.3 RAS-*sr*-DFT dissociation energy curves of H_2 for the ground state singlet (S_0) computed with different short-range DFT functionals (upper left plot), with RAS-*sr*-LDA and different values for the range separation parameter μ (upper right plot). And computed with different *sr*DFT functional and different high-spin triplet reference configurations (ROHF and ROKS) (lower central plot). The Slater exchange functional has been employed for the ROKS reference, while the correlation functional corresponds to the full range counterpart to the one used in the *sr*DFTc calculation, *i.e.*, VWN or PW92. All calculations have been done with the 6-31G basis set. 68
4.4 Ground state dissociation energy curves (in kcal/mol) of H_2 with respect to the dissociation limit computed at the FCI, RAS-SF, RAS-*sr*LDA (*sr*-LDA), RAS-*sr*LDA with with spin polarized densities (*sr*-*p*LDA), with short-range HF exchange (*sr*-EXX), and with short-range correlation (*sr*-LDAc) levels (left plot). And RAS-*sr*-*p*LDA dissociation energy curves of H_2 for the ground state singlet (S_0) and the T_1 with $M_S = \pm 1$ triplet microstates (right plot). Calculations done with the 6-31G basis set. 69

4.5 Relative error in triatomic diradical series. Errors correspond to the ST gaps with respect to the Expt. $-\Delta$ ZPE value reported from references as indicated in previous tables. 74
4.6 Accurate singlet-triplet energy gaps computed for benzyne isomers (in eV). Distribution of the unpaired electrons in the real space is correctly described by the fractional occupation density FOD. 75
4.7 The π-system of TMM at the flat D_{3h} symmetry of the $^3A_2'$ ground state geometry computed at the ROHF/cc-pVTZ level. 77
4.8 Low lying states of TMM biradical and respective FODs and effectively unpaired electrons accounted by the Head-Gordon's N_u index. 78
4.9 ROHF/cc-pVDZ ($^3\Sigma_g^-$) frontier molecular orbitals defining the RAS2 space employed in RAS-SF and RAS-SF-*sr*DFTc calculations of Ga_2. 80
4.10 Dissociation energy curves (in Hartrees) of the ground and low-lying states of the Ga_2 computed at the RAS-SF-*sr*DFTc/6-31G(d) level. Energies are given with respect to the dissociation limit. 81

5.1 Representation of the momentum distribution for the HEG. Two points are highlighted, the first at the Fermi wavevector (k_F) with its associated probability limit approaching from $k > k_F$ ($n_+(r_s) = n(k_F^+, r_s)$) and the second with a momentum $k_C > k_F$ and probability n_C that corresponds to the point from which the dynamic correlation can be accounted by the DF. 87
5.2 Ground state NOONs at the tail of the NO space of octatetraene computed with: (a) the cc-pVQZ basis and different truncations of $\hat{R}$, and (b) RASCI(h, p)/cc-pVXZ with X = D, T and Q. 91
5.3 RASCI(h, p) and RAS-*so*VWN (labelled as *so*VWN with $\nu = 0.001$) energy errors (in eV) with respect to FCI values computed with the cc-pVQZ basis set for the He (a) and Be (b) atomic series. S and G in parenthesis indicate the two considered expressions for $n_+(r_s)$, Eqs. 5.18 and 5.19, respectively. 94
5.4 (a) Total electronic energy profiles (in a.u.), (b) sum of small NOONs, and (c) RAS-*so*VWN (*so*VWN in short) energy (in a.u.) along the H_2 dissociation. . . 95
5.5 (a) Absolute energy profiles along the N_2 dissociation computed at the RASCI (dashed black), RAS-*so*VWN ($\nu = 0.001$, solid red), $\mathcal{E}_0^{\mathrm{WFA},\nu}$ contribution (solid blue), and CIPSI (solid black) computed with the cc-pVDZ basis set. (b) Sum of NOONs below the threshold $\nu = 0.001$ (black) and $E_c^{so,\nu}$ (red) along the N_2 dissociation. 99

6.1 Atomic total energies (in a.u.) computed within the CASSCF+V_{eff} represented as contour plots resulting from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot. The CASSCF energy lies in a vertical line at $\mu = 0$. 109
6.2 Coulomb holes computed for singlet atomic systems at FCI/cc-pVDZ (continuous black line), CASSCF (dashed black line), and CASSCF+V_{eff} with different μ (corresponding ν taken from parametrization with FCI). 112
6.3 H_2 PEC comparison of different methods. This CASSCF+V_{eff} accurate result was obtained within the combination of parameters: $\mu = -0.02$ and $\nu = 0.50$, obtained by computing dissociation energies with different (μ, ν) combinations, and selecting those closest to the FCI result. 113

6.4 H_2 Coulomb holes at equilibrium geometry, computed within CASSCF(2,2), FCI and CASSCF(2,2)+V_{eff} with the cc-pVDZ basis set. The values for the parameters in the effective potential are depicted in the figure. 114
6.5 (a) Most accurate H_2 PECs estimated from parametrization based on the dissociation energies. Effect of different ν values. (b) Corresponding Coulomb holes at equilibrium geometry ($r_{H-H} = 0.8$ Å). In all cases, $\mu = -0.02$ and the cc-pVDZ basis set were used. 115
6.6 Computed Coulomb holes for (a) N_2 and (b) NO^+ diatomics at equilibrium bond lengths. 119
6.7 Ill-behaved PES ($\nu = 1.0, 1.5, 2.0$) for triplet ground state O_2, results computed with larger ν values show considerable distortions to the PES. Percentage of errors with respect to experimental dissociation values follows the same color code as that indicated by labels in the plot. 119
6.8 (left) Obtained H_2 PEC estimated from PS2 (computed with $\nu = 0.50$), the CASSCF+V_{eff} total energy is in perfect agreement with FCI/cc-pVTZ energy. (right) Corresponding optimum μ_{inter} values changing along the separation of H atoms, i.e. those for which CASSCF+V_{eff} energies are lower or equal to 2 kcal/mol. 121
6.9 Optimum μ_{inter} values for O_2 along the PEC (computed with fixed $\mu_{intra} = 0.001$ value). Results computed with ν values: (a) 0.10, (b) 0.50, (c) 1.00, and (d) 1.50. A comparison of these figures reveals hints about the sensitivity of μ_{inter} values to ν parameter. 121
6.10 Illustration of μ_{inter} values which give CASSCF+V_{eff}/cc-pVDZ energies ≤ 2.00 kcal/mol with respect to FCI, along the PEC of N_2 (red), NO^+ (black) and NH (green). Results obtained with fixed $\mu_{intra} = 0.001$ value, and ν values: (a) 0.10, (b) 0.50. Note the pattern distortion increasing with the ν in comparing (a) with (b) cases, also note an enhanced effect present in heteroatomic dimers. 122
6.11 PECs delivering the most accurate D_e for (a) N_2 ($\mu_{\text{intra}} = 0.01$), (b) NO^+ ($\mu_{\text{intra}} = 0.01$), (c) NH ($\mu_{\text{intra}} = 0.00$) and (d) O_2 ($\mu_{\text{intra}} = 0.01$), computed with the erf-fitting strategy (blue continuous lines) a and b values used for the calculations can be consulted in Tables 6.6 and 6.7. The CASSCF/cc-pVTZ (black solid line), NEVPT2/cc-pVTZ correction (black solid line with dots), and CIPSI-FCI/cc-pVTZ (black dashed line) results are included for comparison. 123
6.12 Correlation indexes $\overline{I}_D$ and $\overline{I}_{ND}$ computed at CASSCF/cc-pVDZ level, for the four diatomics (a) N_2, (b) NO^+, (c) NH, and (d) O_2 along the dissociation coordinate. 127
6.13 PECs delivering the most accurate D_e for (a) N_2, (b) NO^+, (c) NH and (d) O_2, computed with correlation indices fitting (red continuous line). Optimum fitting X, Y values used for the calculation can be consulted in Table 6.8. The CASSCF/cc-pVTZ (black solid line), NEVPT2/cc-pVTZ correction (black solid line with dots), and CIPSI-FCI/cc-pVTZ (black dashed line) results are included for comparison. 128

7.1 Representation of three investigated structural configurations of the TM biradical: (0,0) and (90,90) with C_{2v} symmetry, and the C_s structure (0, 90). 138
7.2 Dominant NTO pairs (isovalues of 0.05) of the spinless transition matrix of TM involved in the singlet-triplet SOC at (0, 90) and (90, 90) torsion angle configurations. The respective singular values ω_1 and ω_2 are depicted below each pair. 139

7.3 Triplet state orbital diagram corresponding to the RAS2 space (8e, 7o) of planar ethylene computed at the ROHF/cc-pVTZ level. 141
7.4 Singlet-triplet total and 1-electron RAS-SF/cc-pVTZ (black) and 1-electron CIS/STO-3G (blue) SOCCs, and $I_C = c_H + c_L$ and $I_U = (2 - N_U)/2$ indices (red) along the twist modes of ethylene. N_U correspond to the Head-Gordon index for the singlet state. 143
7.5 Dissociation curves for ethylene at 0° (a) and 50° (b) twist configurations computed at the RAS-SF/cc-pVQZ level. The graphs show the five lowest energy states selected for analysis. 144
7.6 Yamaguchi (N_Y) and Head-Gordon (N_U) indexes for lowest singlet state along the dissociation of ethylene. The three discussed regions are indicated as CS (closed-shell), D (diradical), T (tetraradical). 145
7.7 (a) SOCC profiles (in cm^{-1}, eq 7.17) between TT and ST manifolds of ethylene at 0° (black solid line) and 50° (red solid line) twist configurations along the C=C dissociation (RAS-SF/cc-pVQZ). Dashed horitzontal lines indicate the ${}^3B_2/{}^1A_1$ coupling in the CH_2 fragment. (b) Squared amplitudes of the main configurations for the ${}^3B_{2g}$ state along the stretching coordinate. 146
7.8 Plot of SOCCs values for ethylene at 0° twist configuration along the C=C stretching from equilibrium distance (RAS-SF/cc-pVQZ). 146
7.9 SOCC profiles for ethylene at 0° twist configuration along the C=C stretching from equilibrium distance (RAS-SF/cc-pVQZ). 147
7.10 Two dominant NTO pairs of the spinless transition matrix (isovalue of 0.05) between ${}^1A_g/{}^3B_{2g}$ (a) and ${}^3B_{1u}/{}^3B_{3u}$ (b) states of planar ethylene computed for different CC distances at the RAS-SF/cc-pVQZ level. The two largest singular values are indicated as ω_1 and ω_2, respectively. 148
7.11 First dominant NTO pairs of the spinless transition matrix of ethylene (RAS-SF/cc-pVQZ) transitions at different stretch distances (isovalues of 0.05). The two largest singular values (ω_1 and ω_2) between the NTO pairs are presented below each pair. 149
7.12 Ethylene-oxygen dimer structures with the representation of the FOD of the ground state triplet (blue isosurface). 151
7.13 (a) Energy profiles (in eV) of low-lying states, and (b) SOCC (in cm^{-1}) between X^3B and the two lowest A-symmetry singlets, ${}^1A(\Delta)$ and ${}^1A(\mathrm{CT})$ states, along the coplanar torsion between **eclipsed** and **crossed** conformers computed at the RASCI/cc-pVQZ level of theory. Energies are given with respect to the ground state triplet at the **crossed** structure. 153
7.14 Dominant NTO pairs (isovalues of 0.05) of the spinless transition matrix between X^3B and the two lowest A-symmetry singlets for the $C_2H_4 \cdots O_2$ system with a 55° torsion angle. The participation ratios (PR) and the two largest singular values (ω_1 and ω_2) for the NTO pairs are also indicated. 154
7.15 Energy (in eV) and SOCC (in cm^{-1}) profiles for the lowest lying X^3B and ${}^1A$ states of the $C_2H_4 \cdots O_2$ system at 75° (a and b) and 15° (c and d) torsion angles computed at different intermolecular separations. 155

1 Number of unpaired electrons for the H_2 ground state along the molecular dissociation computed at the FCI, RAS-SF, RAS-*sr*LDA (*sr*-LDA), RAS-*sr*LDA with with spin polarized densities (*sr*-*p*LDA), with short-range HF exchange (*sr*-EXX), and with short-range correlation (*sr*-LDAc) levels and the 6-31G basis set. 161

2 Illustrative NOONs computed at the RASCI(h,p)/cc-pVQZ level for $C_{2n}H_{2n+1}$ linear polyenes (n = 2-4). 163
3 RASCI(h,p)/cc-pVDZ natural orbital occupation numbers (NOONs) representation for the isoelectronic He-series. Each horizontal line represents a natural orbital from the active space (RAS2(2e, 5o)), the numerical value of the respective NOON is placed above. Reported calculations where performed within this selection of the active space. 165
4 Full valence RASCI(h,p)/cc-pVDZ NOONs representation for the isoelectronic Be-series. Each horizontal line represents a NO from the active space (RAS2(4e, 5o)), the numerical value of the respective NOON is placed above. 165
5 Correlation energies computed for the Be-Series with: *on-top* RAS(NOF) and RAS-NOF schemes. NOONs ($\nu_{threshold} = 0.001$) was chosen in order to get a comparison with the results reported with $\nu_{threshold} = 0.01$. Computed energies are close to the FCI energy than the previous case. 167
6 Correlation energies computed for the Be-Series within the RAS-NOF scheme $\nu_{threshold} = 0.01$. The use of a large basis set improves the calculation, however a similar overestimation as that obtained with the cc-pVDZ basis set remains. 167
7 Complete data for (a) sum of small NOONs, and (b) soVWN energy (in a.u.) along the H_2 dissociation. 169
8 H_2 PES comparison among the $on-top$ (a and b) and the NO splitting (c and d) schemes, threshold values are indicated in the labels. 170
9 DFT energy dependence on the parameter nu for the H_2 ground state at $r_{HH} =$ 0.74 Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets. 173
10 Total energy dependence on the parameter nu for the H_2 ground state at $r_{HH} =$ 0.74 Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets. 173
11 Ethylene potential energy surfaces for the singlet ground state and the lowest lying triplet state. Left: RAS-soDFT profiles obtained with the Savin (left) and Gori (right) $n_+(r_s)$ expressions, and with different values for the sum of occupations threshold. 174
12 DFT energy dependence on the parameter nu for the C_2H_4 ground state (optimized at the CCSD(T)/cc-pVDZ level) computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets. 177
13 Total energy dependence on the parameter nu for the C_2H_4 ground state (optimized at the CCSD(T)/cc-pVDZ level) computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets. 177
14 Computed values for the effective unpaired electrons accounted from the Head-Gordon "nu" and "nu_nl" indexes along the dissociation of N_2. 178
15 Natural orbital occupations (higher than a 1.0E-05) at different separations. . . 179
16 Natural orbital occupations (lower than a threshold of 0.001) at different separations. 180
17 Absolute energy profiles along the N_2 dissociation computed at the RASCI (dashed black) and CIPSI (solid black) computed with the cc-pVDZ basis set and their comparison with (a) *on-top* DFT calculations and (b) splitting the natural orbital space. In both cases the 1S single NOON based threshold scheme was used (number in label resembles the occupation threshold used). 181

18 General comparison of N_2 PES computed with the different schemes presented in this work and standard RASCI and CIPSI calculations within the cc-pVDZ basis set. The $E_>$ label indicates such cases where no DFT correlation energy is incorpored into the final energy, i.e. these energies correspond to the WFT energy computed in the space of NOs with large occupations. 181
19 DFT energy dependence on the parameter nu for the N_2 ground state at $r_{NN} =$ 1.10 Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets. 187
20 Total energy dependence on the parameter nu for the N_2 ground state at $r_{NN} =$ 1.10 Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets. 188
21 Absolute energy profiles along the N_2 dissociation computed at the RASCI (dark blue), RAS-soVWN (yellow), RAS-soVWN (orange) computed with the cc-pVDZ basis set. RAS-soVWN have been obtained by fixing the number of NOs entering the soVWN correlation to those corresponding whose sum of occupation numbers is lower than 0.001 at the equilibrium distance. 188
22 Sum of occupations lower than 0.001 and corresponding values of the DFT correlation functional soVWN(G). Results computed with different basis sets. . 189
23 Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2 \cdots$ He system. 190
24 The different $H_2 \cdots n$He systems considered in this analysis. He atoms are placed orthogonally in the x- and y-axes, while H_2 bond lies on the z-axis. . . 190

25 B, C and N atoms total (left plots) and correlation (right plots) energies (in a.u.) computed within the CASSCF+V_{eff} represented as contour plots resulting from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot. 208
26 O, F and Ne total (left plots) and correlation (right plots) energies (in a.u.) computed within the CASSCF+V_{eff} represented as contour plots resulting from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot. 209
27 C^+, N^+, O^+ cations CASSCF+V_{eff} obtained contour plots of total (left plots) and correlation energies (right plots), from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot. 210
28 F^+, Ne^+, cations CASSCF+V_{eff} obtained contour plots of total (left plots) and correlation energies (right plots), from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot. 211
29 Anions CASSCF+V_{eff} obtained contour plots of total (left plots) and correlation energies (right plots), from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot. 212
30 Coulomb holes for the C atom with fixed $\nu = 0.01$ value and variations of μ (both positive and negative), computed with the cc-pVDZ basis set. 213
31 Evaluation of the effect of a positive effective potential. Coulomb holes for the C atom with fixed $\nu = 0.09$ value and variations of μ, computed with the cc-pVDZ basis set. 214
32 Evaluation of the effect of a positive effective potential. Coulomb holes for the C atom with fixed $\mu = 0.02$ value and variations of ν. 214

33 H_2 PECs computed with different values of μ (positive and negative), small variations lead to considerable changes in the PES, $\nu = 0.09$. Here the use of $\mu = +0.03$ leads to apparent distortions at long-range regions in the PEC (dashed blue line). 215
34 (Up) Potential energy surfaces of H_2 computed with selected parameters yielding energles close to FCI/cc-pVDZ. Changes in ν do not lead to large variations of the PES. (Down) Corresponding Coulomb holes at equilibrium. Small varitions do not deteriorate the Coulomb hole with respect to CASSCF. 216
35 Best estimates computed for N_2 PECs, giving dissociation energies close to experimental values taken from Ref. [2] (calculations performed with $\nu = 0.5$ fixed value). 217
36 Best estimates computed for NO^+ PECs, giving dissociation energies close to experimental values taken from Ref. [2] (calculations performed with $\nu = 0.5$ fixed value). 217
37 The lowest value considered $\nu = 0.1$ gives the best PES for O_2. Percentage of errors with respect to experimental dissociation values (taken from Ref. [2])follows the same color code as that indicated by labels in the plot. 218
38 Selected PES for NH case, labels show the values of the two μ parameters in the effective potential together with the ν selection. Percentage of errors with respect to experimental dissociation values (from Ref. [2]) follows the same color code as that indicated by labels in the plot. 219
39 Selected PES for OH doublet, results computed with $\nu = 0.1$. As in previous examples, the dissociation errors are depicted withing the same color code as the plot. 220
40 Selected PES for OH doublet, results computed with different values of ν exponential parameter, small changes to the PES are obtained for the different combinations of the multiplicative μ parameters. As in previous examples, the dissociation errors are depicted withing the same color code as the plot. 220
41 Fitted parameters and comparison with a fraction of the atomic volumes (V_A/100 and V_B/100) and delocalization index (D(A,B)) computed at equilibrium. Results computed with different ν values as indicated, $\mu_{intra} = 0.000$ (for (a) and (b) cases) and $\mu_{intra} = 0.001$ (for (c) and (d) cases). 221

42 Comparison of results, with cc-pVQZ basis set, among the SOCC, the ST gap and the SOC_{rms} for TM. The energy scale is shown only for illustrative purposes, SOCCs are in cm^{-1} while the values of the vertical S-T gaps are presented as ten times their original value (in eV). Similar qualitative trends are found by the two approaches. 227
43 Dominant NTO pairs of the spinless transition matrix of ethylene (ROHF/cc-pVTZ triplet as reference) in the $T_1 \leftarrow S_0$ transitions at different torsion angles (isovalues of 0.05). The two largest singular values (ω_1 and ω_2) and participation ratios (PR) are presented on the right. 228
44 Reference ROHF quintet orbitals (which forms RAS2) for the D_{2d} configuration, and its evolution along the C=C stretching. 228
45 Reference ROHF quintet orbitals (which forms RAS2) for the D_2 configuration, and its evolution along the C=C stretching. 229

46 Dissociation curves for ethylene at 0° twist configuration computed with cc-pVQZ basis set. The graph shows the twelve selected states, those with lines in blue (and violet) are the ones that have non-zero spin orbit coupling with the three lowest lying states at dissociation 229
47 Dissociation curves for ethylene at 50° twist configuration computed with cc-pVQZ basis set. The graph shows the twelve selected states, those with lines in blue are the ones that have non-zero spin orbit coupling with the three lowest lying states at dissociation. The only state that doesn't couples is the triplet ^3A state. 230
48 SOCC profiles for ethylene at (a) 0° and (b, c) 50° twist configurations along the C=C separation (RAS-SF/cc-pVQZ). The value of the MF-SOCC for the $^3B_2/^1A_1$ coupling from the CH_2 fragment is represented by the black-dashed line, the double of this quantity is indicated by the black-dotted line. Total SOCC is represented by the red curve. 231
49 Plot of squared amplitudes values for ethylene's $^3B_{3u}$ state along C=C separation (RAS-SF/cc-pVQZ). 232
50 First dominant NTO pairs of the spinless transition matrix of ethylene (RAS-SF/cc-pVQZ) transitions at different stretch distances (isovalues of 0.05). The two largest singular values (ω_1 and ω_2) between the NTO pairs are presented below each pair. 233
51 Dominant NTO pairs of the spinless transition matrix of $C_2H_4\cdots O_2$ configuration involved in the $^3B \leftrightarrow ^1A$ and $^3B \leftrightarrow 3^1A$ transitions near the crossing regions, at 2.4 Å (15° rotation) and 2.1 Å (75° rotation) separations of O_2 fragment, with isovalues of 0.05. The participation ratios (PRts) and the two largest singular values (ω_1 and ω_2) between the NTO pairs are presented below each case. 236

List of Tables

4.1 Energy gaps (in eV) between the ground state triplet (^3P) and the lowest singlet state (^1D) in carbon and silicon atoms, computed with RAS-SF (high-spin quintet as the reference configuration) and RAS-IP (high-spin quartet anion as the reference configuration) with the pristine wave function (RASCI) and different implementations of RAS-*sr*DFT and the cc-pVTZ basis set. ΔE_0 and ΔE_1 values correspond to energy gaps with respect to the $M_S = 0$ and $M_S = \pm 1$ ^3P triplet, respectively. Experimental gaps indicated in parenthesis. 71

4.2 Energy gaps (in eV) between the ground state triplet (^3P) and the lowest singlet state (^1D) in oxygen and sulfur atoms, computed with RAS-IP (singlet dianion as the reference configuration) and RAS-AE (high-spin quartet cation as the reference configuration) with the pristine wave function (RASCI) and different implementations of RAS-*sr*DFT and the cc-pVTZ basis set. ΔE_0 and ΔE_1 values correspond to energy gaps with respect to the $M_S = 0$ and $M_S = \pm 1$ ^3P triplet, respectively. Experimental gaps indicated in parenthesis. 71

4.3 Adiabatic energy gaps (in eV) with respect to the $\widetilde{X}^3B_1$ ($M_s = 1$) ground state for CH_2 and NH_2^+ molecules computed with RAS(n, m)-SF and RAS(n, m)-SF-*sr*LDAc, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively. 72

4.4 Adiabatic ST gaps (in eV) with respect to the $\widetilde{X}^1A_1$ ground state for SiH_2 and PH_2^+ molecules computed with RAS(n, m)-SF and RAS(n, m)-SF-*sr*LDAc, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively. 73

4.5 Singlet-triplet adiabatic energies (in eV) of benzyne isomers computed with RAS-SF, RAS-*sr*DFT and RAS-*sr*DFTc with two different RAS2 spaces, and with the ROHF (HF superindex in parenthesis) and ROKS (KS superindex in parenthesis) triplet reference configurations. Spin projection of the triplet state indicated with a subindex 0 ($M_S = 0$) or 1 ($M_S = \pm 1$). 76

4.6 Vertical excitation energies (in eV) to the low-lying excited states with respect to the ground $^3A_2'$ state of TMM computed at the RAS-SF and RAS-SF-*sr*LDAc (RAS-*sr*LDAc) with the cc-pVTZ basis set. The (n, m) values indicate the number of electrons (n) and orbitals (m) in the RAS2 space. 78

4.7 Binding energies D_e for the Ga_2 $^3\Pi_u$ ground state (in eV) and adiabatic excitation energies (in kJ/mol) computed with RAS-SF and RAS-SF-*sr*DFTc (*sr*DFTc in short) for the low-lying excited states with respect to the ground state. 81

5.1 Dissociation energies (in mHartrees) for H_2 within the cc-pVDZ basis set. These results illustrate three different partitions of the NO space defined through the values for the threshold. $\mathcal{E}_0^{\mathrm{WFA},\nu}$ wave function contribution defined in Eq. 5.11. 96

5.2 Energy barrier (in kcal/mol) computed for the molecular torsion of ethylene. RAS-soVWN (indicated as soVWN), with values in parenthesis corresponding to the employed sum of occupations threshold. 97
5.3 N_2 dissociation energies (in kcal/mol) computed with CASSCF, RASCI, $\mathcal{E}_0^{\text{WFA},\nu}$, RAS-$so$VWN ($so$VWN in short) with the cc-pVDZ, and compared to reference results from the literature. 98

6.1 Full valence CASSCF, FCI, and CASSCF+V_{eff} ($\mu = -0.006$) correlation energies for first row atoms and ions (in a.u.). Calculations performed with the cc-pVDZ basis set. Corresponding ν values computed from the data fitting ($r^2 > 0.99$). 110
6.2 Atomic ground state ionization potentials and electron affinities in eV. 111
6.3 Most accurate CASSCF+V_{eff} ($\mu = -0.02$) dissociation energies and percentage of error with respect to FCI/cc-pVDZ dissociation energy for H_2 system. . . . 114
6.4 Effective potential parameters giving the lowest percentage of error in the computed dissociation energy with respect to FCI. 118
6.5 Effective potential parameters giving the lowest percentage of error in the computed dissociation energy with respect to FCI. 120
6.6 Singlet state diatomics, resulting dissociation energies (in eV) and percentage of errors with respect to FCI from calculations with μ_{inter} fitted to the function $a(\text{erf}(br_{AB} - 1))$. 124
6.7 Triplet state diatomics, resulting dissociation energies (in eV) and percentage of errors with respect to FCI from calculations with μ_{inter} fitted to the function $a(\text{erf}(br_{AB} - 1))$. 125
6.8 Resulting dissociation energies and percentage of errors with respect to FCI from calculations optimizing $\mu_{inter} = -0.1I_D$ and using different choices for $\mu_{intra} = YI_{ND}$. 129

7.1 RAS-SF SOCCs (in cm^{-1}) computed with (6e, 5o) RAS2 space and the cc-pVXZ (X = D, T, and Q) basis set family. Values for the Head-Gordon N_U index for the singlet state and vertical singlet-triplet gap ($\Delta E_{ST} = E_S - E_T$ in meV) computed at the RAS-SF/cc-pVQZ level. 138
7.2 Singlet-triplet SOCC (in cm^{-1}) computed for different TM conformers with CASSCF and RAS-SF with different truncation of the excitation operator. All values obtained with an active space (RAS2) with 2 electrons in 2 orbitals and with the 3-21G basis set. 139
7.3 Singlet-triplet SOCC (in cm^{-1}) computed for different TM conformers at the RAS-SF/cc-pVQZ level with three different RAS2 spaces. 140
7.4 Systematic comparison of SOCCs (in cm^{-1}), effect of the RAS2 active space indicated in parenthesis computed with the 6-31G basis set. 140
7.5 Singular values for the leading NTO pairs and participation ratios of the natural transition orbitals (PRNTOs) for the 0° and 50° twist configuration along dissociation, the equilibrium C=C distance for the planar (50° twist) geometry is 1.3516 (1.4794) Å. Here C=C (Å) regards to the stretch distance from equilibrium. 150
7.6 SOCC (in cm^{-1}) and vertical energy gaps (in eV) between the lowest triplet and singlet states of the ethylene-oxygen dimer structures in Figure 7.12 computed at the RAS-SF/cc-pVQZ level. 152

1 Adiabatic energy gaps (in eV) with respect to the $\widetilde{X}^3B_1$ ($M_s = 1$) ground state for CH_2 and NH_2^+ molecules computed with RAS(n, m)-SF-*sr*LDA, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively. 162
2 Adiabatic ST gaps (in eV) with respect to the $\widetilde{X}^1A_1$ ground state for SiH_2 and PH_2^+ molecules computed with RAS(n, m)-SF-*sr*LDA, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively. 162

3 Sum of NO occupations for the ground singlet state ($1^1A_g^-$) in each RAS subspace for linear n-polyenes ($C_{2n}H_{2n+1}$). 164
4 Sum of NO occupations for the ground singlet state ($1^1A_g^-$) in each RAS subspace for linear n-polyenes ($C_{2n}H_{2n+1}$). Sums for the RASCI(p) and the RASCI(h,p) are the same, as for the RASCI(h) and CASCI calculations. Particle contributions are determinant for the presence of fractional occupancies at RAS3 space. . . . 164
5 [a]The 'exact' values of the correlation energies (E_c^{Exact}) were taken from Reference [3], and correspond to the difference between corrected experimental values and Hartree-Fock energies. A RAS2(2e, 5o) and RAS2(4e,5o) active spaces were used for the He- and Be-series respectively. An occupation threshold of 0.001 was chosen. 166
6 Total energies (in a.u.) for the He-series computed within different approaches as discussed in main text (with the soVWN(G) functional and $\nu = 0.001$), the cc-pVQZ basis set was used. 168
7 Total energies (in a.u.) for the He-series computed within different approaches as discussed in main text (with the soVWN(S) functional and $\nu = 0.001$), the cc-pVQZ basis set was used. 168
8 Total energies (in a.u.) for the Be-series computed within different approaches as discussed in main text (with the soVWN(G) functional and $\nu = 0.001$), the cc-pVQZ basis set was used. 168
9 Total energies (in a.u.) for the Be-series computed within different approaches as discussed in main text (with the soVWN(S) functional and $\nu = 0.001$), the cc-pVQZ basis set was used. 169
10 Dissociation energies (in mHatree) for H_2 computed with different basis sets. Threshold for the sum of occupations used is 0.001. Experimental value is 164.5 mHartree.[4] 170
11 Corresponding total energies for the H_2 PES computed in Figure 8 within the cc-pVDZ basis set in the *on-top* approach. The threshold value is indicated inside squared brackets below the RAS-soVWN functional labels. 171
12 Corresponding total energies for the H_2 PES computed in Figure 8 within the cc-pVDZ basis set in the RAS-soDFT approach. The threshold value is indicated inside squared brackets below the RAS-soVWN functional labels. 172
13 PES data for the ethylene singlet state torsion (angle in degrees) corresponding values from Figure 11. 175
14 PES data for the ethylene triplet state torsion (angle in degrees) corresponding values from Figure 11. 176

15 Correlation energies (computed with soVWN(S) functional) and E_ν values computed within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold. 182
16 RAS-soVWN(S) total energies and comparison with CASSCF and FCI calculations within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold. Dotted empty values cannot be computed since the threshold values pushes to cut the RAS2 space. 183
17 Correlation energies (computed with soVWN(G) functional) and E_ν values computed within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold. 184
18 RAS-soVWN(G) total energies and comparison with CASSCF and FCI calculations within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold. Dotted empty values cannot be computed since the threshold values pushes to cut the RAS2 space. 185
19 *On-Top* PES total energy results computed within the soVWN(S) and soVWN(G) functionals, the cc-pVDZ basis set and different threshold values as indicated. 186
20 Selected RAS-soDFT results PES total energy results computed within the soVWN(S) and soVWN(G) functionals (the threshold value was set to 0.001 in both cases), the cc-pVTZ basis set and different threshold values as indicated. . 187
21 Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2\cdots He$ system. 190
22 Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2\cdots 2\,He$ system. 191
23 Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2\cdots 3\,He$ system. 191
24 Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2\cdots 4\,He$ system. 191

25 Optimal parameters for atomic systems computed from a linear fitting from computed data (variations of μ and ν parameters). 213
26 Full table with optimal fitting values to the function $a(erf(bR_{AB}-1))$, R^2 value and their comparison with atomic volumes (V_A and V_B) and delocalization index DI(A,B) obtained as indicated in main text. 222

27 SOC components and constants (in cm^{-1}) for the different geometries of TM. Calculations performed with different basis sets and a (6e,5o) RAS2 space. . . . 225
28 Continuation, SOC components and constants (in cm^{-1}) for the different geometries of TM. Calculations performed with different basis sets and a (6e,5o) RAS2 space 226
29 Systematic comparison of SOCCs (in cm^{-1}), effect of the RAS2 active space indicated in parenthesis computed with the 6-31G basis set. 226
30 Selected states and dominant electronic configurations along dissociation at 0° and 50° torsion cases. 232
31 Singular values and participation rations for the leading NTO pairs for the non-analogous couplings that appear at 50° twist configuration. 234
32 Singular values and participation rations for the leading NTO pairs couplings with higher energy states. 234
33 Leading configurations and Head Gordon's N_u values at 5° and 55° of O_2 rotation (2.00 Å of separation), for higher energy states in order of magnitude of the RASCI weight. 235
34 Vertical excitation energies (in eV), of the lowest energy states for the EtO1 system under rotation of the O_2 fragment (RAS-SF/cc-pVQZ calculations). Only illustrative data was chosen. 235
35 Leading configurations and Head Gordon's N_u values at 5° and 55° of O_2 rotation (2.00 Å of separation), for the target states in order of magnitude of the RASCI weight. 236

CHAPTER 1

Introduction

Theoretical chemistry is devoted to the study, elucidation and understanding of chemical phenomena through the use of electronic structure theory methods, whose foundations are well established and understood. For a molecular system of N electrons and M nuclei described by the position vectors $\vec{\boldsymbol{r_i}}$ and $\vec{\boldsymbol{R_A}}$ respectively, the full non-relativistic Quantum Mechanical Hamiltonian can be written in atomic units as

$$\hat{H} = -\sum_{i}^{N} \frac{1}{2}\nabla_i^2 - \sum_{A}^{M} \frac{1}{2M_A}\nabla_A^2 - \sum_{i}^{N}\sum_{A}^{M} \frac{Z_A}{r_{iA}} + \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} + \frac{1}{2}\sum_{A}^{M}\sum_{B\neq A}^{M} \frac{Z_A Z_B}{R_{AB}}, \tag{1.1}$$

where the first and second terms are the kinetic energy terms for electrons and nuclei, respectively. Labels A, B, ..., are for nuclei and i, j, ..., are for electrons, Z_A (Z_B) are the atomic number of the nucleus A (B) and M_A is the ratio of the mass of nucleus A to the mass of an electron. The distance between electron i and nucleus A is $r_{iA} = |\vec{\boldsymbol{r_i}} - \vec{\boldsymbol{R_A}}|$; the distance between electron i and electron j is $r_{ij} = |\vec{\boldsymbol{r_i}} - \vec{\boldsymbol{r_j}}|$ and the distance between nuclei A and B is $R_{AB} = |\vec{\boldsymbol{R_A}} - \vec{\boldsymbol{R_B}}|$. The Coulomb attraction term between the ith-electron and the Ath-nucleus is $1/r_{iA}$, the corresponding electron-electron (nucleus-nucleus) interaction term is $1/r_{ij}$ ($1/R_{AB}$).

The electronic structure description of ground and excited states is often based on the Born-Oppenheimer approximation. Within this approximation, the nuclear kinetic energy term is neglected and the last nucleus-nucleus repulsion term in 1.1- can be considered constant. Consequently, the complex many-body quantum problem of treating nuclei and electrons in a molecular system is reduced to find the solution to the time-independent electronic Schrödinger equation (eSE):

$$\hat{H}_{el}(r;R)\Psi(r;R) = E(R)\Psi(r;R), \tag{1.2}$$

where $\hat{H}_{el}(r;R)$ is the electronic Hamiltonian, $\Psi(r;R)$ the electronic wave function that depends on the coordinates of each electron of the system (r) and parametrically on the positions of the nuclei (R). The total electronic energy $E(R)$ is still a function that depends on the nuclear spatial configuration for a given molecular system. The main goal of electronic

structure methods is to get an accurate (close to or nearly exact) solution for Eq. 1.2. In fact, a variational (exact) solution exists, the so-called full configuration interaction (FCI), however, due to the factorial scaling growth of the computational cost needed to acquire such solution, its application is currently restricted to systems of no more than 20 electrons [5, 6, 7]. Over time, this constraint has motivated the objective of developing approximate but well-defined mathematical procedures of simulation, and devising novel methods capable of diminishing the computational cost for accurate solutions of the eSE [1]. Approximations based directly on the eSE, are commonly referred as wave mechanics of wave function theory (WFT). Alternatively, a powerful electronic structure method, based on the electronic density rather than electronic wave functions is called density functional theory (DFT) [8, 9], the most popular version of which is based on the Kohn-Sham (KS) equation.

In any case, achieving accurate solutions with an affordable computational cost involves addressing inherent challenges, such as:

(i) The formidable difficulties posed by the eSE, being a non-separable second-order partial differential equation in $3 \times N$-spatial coordinates (with N being the number of electrons). The eSE has not only one, but infinitely many solutions that are antisymmetric under the interchange of any two electrons. This illustrates how difficult the task of solving the equation is.

(ii) In practice, for approximate solutions, numerical precision challenges in connecting the eSE's energies to properties of interest become important. This is due to the inherent scaling of such energies with the molecular size (extensive), but the magnitudes of intensive energies (such as bond energies, activation energies, excitation energies, ionization potentials, and electron affinities, etc.) do not change correspondingly. Therefore, to reliably compute the intensive properties by subtracting two total energies (computed from eSE) requires the extensive energies be computed to high precision as the molecular size grows. While some methods have been designed to account for this problem by expressing intensive properties as solutions to matrix eigenvalue problems, as embodied, for example, in equations-of-motion (EOM) or Green's function methods [10, 11, 12], there are properties such as reaction activation energies and non-covalent intermolecular interactions that still need to be determined through eSE energy differences (see, for example Ref. [13]).

(iii) Orbitals and antisymmetric wave functions (e.g., Slater determinants) give rise to large (third power and higher) scaling in computational effort as the number of electrons and basis functions increase [14]. This is important since, in most cases, large basis sets are required in order to achieve the desired accuracy.

At a fundamental level, a critical factor in refining electronic structure calculations is the treatment of the quantum mechanical Coulomb electron-electron interaction. This is known as the electron correlation problem, and represents a crucial challenge that has garnered significant attention over the years. In the early days of quantum chemistry, the neglect of correlation was identified as leading to severe problems such as the energy overestimation of bond dissociation energies [15, 16, 17]. This deficiency becomes apparent when examining the process of complete homolytic dissociation of a bond, as in the case of H_2. Without correlated motion between the electrons, there exists a finite possibility of both electrons being on the same atomic center [1].

This lack of electron correlation is implicit in single Slater determinant wave functions. Improved wave functions incorporate, somehow, the effect of excited configurations (determinants). In practice, this is commonly performed by expanding the Hartree-Fock (HF)

determinant through perturbation theory, configuration interaction (CI) or through coupled-cluster (CC) ansätze. In KS-DFT the problem is translated to the definition of the exchange correlation functional. However, as pointed out in a recent work [18], the DFT density can be improved to FCI quality through the utilization of a large basis set. Moreover, in the limit of a complete basis set, a reference electron density of the FCI type can be accurately reproduced by a set of orbitals in a single Slater determinant.

Including electron correlation has a significant effect on calculated energies, with implications in a numerous amount of physical and chemical phenomena [15, 19, 20]. To name a few examples: the calculation of bond energies, bond dissociation, potential energy surfaces of main-group and transition metal compounds, reaction barrier heights, proton affinities, hydrogen bond energies, ground- and excited-state charge transfer, valence and Rydberg excitations, and singlet-triplet splitting in diradicals stand out as some of the most important properties of interest. In this regard, WFT has a great disadvantage, since, computational resources grow considerably with system size. A more favorable scaling is achieved by DFT; for this reason, it has become the method of choice for large and complex systems.

The respective advantages and limitations of WFT and DFT have inspired the design of mixed approaches. For small enough systems, WFT can be made systematically accurate, so, there is a motivation to incorporate systematic WFT procedures into DFT to improve the treatment of electron-electron interactions. Mingled WFT+DFT methods attempt to overcome the limitations encountered in both approaches, while taking advantage of their strengths.

This thesis addresses the development and implementation of quantum chemistry hybrid methodologies designed to account for a balanced description of electron correlation within an accessible computational cost, and investigates their performance and limitations. While challenges (i), (ii) and (iii) depicted above represent extremely relevant issues, the problem of electron correlation is central in the development of methods and description of model systems studied in this thesis.

1.1 Overview of this thesis

One of the main objectives of this work is intended to combine WFT methodologies with DFT, resulting in mingled WFT-DFT approximations. We further explore the use of a two-electron effective potential on-top of a WFT method. The rest of this chapter covers a brief introduction to the electron correlation problem, and various definitions of electron correlation energy are introduced. Next, some of the standard *in-use* methods for treating electron correlation are briefly discussed, underscoring the essential criteria that any electronic structure method must consider for its applicability in computational chemistry. Starting from the essentials of HF theory, Chapter 2 is devoted to show the theoretical foundations of electronic structure methods typically encountered in the theoretical chemistry literature. Then a review of active space configuration interaction theories, which, through systematic improvements of the wave function, allows for enhancements in the treatment of electronic correlation. These theories are a basic component for the methods proposed in this thesis. Each of the following chapters contains a short introduction to make them self-contained.

In Chapter 3, we study the capabilities and limitations of WFT combined with short-range DFT (WFT-*sr*DFT) method, and we apply it to the characterization of open-shell systems. Spin polarization effects on the DFT exchange contributions are recognized as the major source of success (or failure) of the method. To this end, we introduce different strategies

to account for spin polarization in the short-range exchange functional. Examples with several model molecular systems allow us to identify a successful strategy for the study of open-shell molecules.

The promising results obtained with the mingled WFT-DFT methodologies is the principal motivation that drives the proposal, development and exploration of the capabilities of a similar approach in Chapter 4. This new method is based on the partitioning of the natural orbital (NO) space, and further discrimination of the available and relevant NO occupations. The idea originates from the experience showing that NOs with large occupations describe relevant, molecule-specific effects. We propose the splitting of the NO space, and introduce, as in previous works with model Hamiltonians [21], a local density functional accounting for the electron correlation associated to NOs with small occupations. We follow the WFT+DFT proposal previously explored by Savin [22, 23], where separation of the NO space is done by means of a given threshold parameter. Orbitals with NO occupation numbers (NOONs) larger than a given threshold (ν) are correlated explicitly at the WFT level, while the remaining correlation energy is described with the complementary ν-dependent NO functional (NOF).

In Chapter 5, we expand the idea previously proposed by Roos [24] on the so-called "correlation factor methods", early mentioned in the review work of McWeeny [25]. In turn, these methods are based on the idea that the short-range correlation (mostly dynamic) remains independent of the CI expansion, but is dealt analytically by introducing interelectronic variables. In fact, the explicit inclusion of the interelectronic separation r_{12} in a 2-particle wave function can lead to solutions of extraordinary accuracy [26, 27, 28, 29], as shown in the calculation of the helium atom by Pekeris [30] and those for the hydrogen molecule by Kolos and Woliniewicz [31]. In our case, the method is based on the combination of an effective two-electron potential that adds dynamic correlation to a wave function including nondynamic correlation, i.e., the complete active state self-consistent field (CASSCF) method. Through an extensive parametrization, and Coulomb hole analysis, we provide both theoretical evidence and numerical proofs, demonstrating the successful application of this strategy in both atomic and diatomic systems.

Finally, Chapter 6 is devoted to show the applicability of the restricted active space configuration interaction (RASCI) method in conjunction with a recent implementation based on the spin-orbit mean field approximation (SOMF), allowing us to explore and unravel the spin-orbit couplings (SOCs) in open shell molecules. In this study, the proper description of nondynamical correlation and the performance of SOMF theory for open-shell systems are analyzed.

Additional information regarding complementary definitions, results and derivations used in the works presented in this thesis can be found in the accompanying appendices.

1.2 Electron Correlation

During the years 1928 to 1929, in the early stages of quantum mechanics, Heisenberg and Dirac [32, 33] described the effects of electron correlation as the fundamental origin of magnetism, specifically in terms of the exchange energy, which depends on the spin state and the wave function's antisymmetry, governed by Pauli's exclusion principle [34]. As previously mentioned, exchange effects appear on electrons of the same spin, keeping them spatially separated. This is a purely quantum mechanical effect with no classical analogue. Despite this initial description, the term was not introduced into quantum chemistry literature until 1934 by Wigner and Seitz as a result of their investigations on the cohesive

energy of metallic sodium [35]. While electron correlation refers to the measure of how the movement of one electron is influenced by the presence of others, the term “electron correlation energy” (E_c) was introduced in 1959 Löwdin, as the difference between the exact eigenvalue of the nonrelativistic electronic Hamiltonian and the HF energy (E_0) computed in the limit that the basis set approaches completeness [36, 37],

$$E_c = E_{exact} - E_0. \tag{1.3}$$

Despite the allowance of exchange effects by the HF wave function, it is inadequate to treat the total correlation among the motions of distinct electrons. The electron correlation energy is a measure of this inadequacy and not an observable linked to an operator. This purely theoretical definition (1.3), was initially introduced by Löwdin [36], and subsequently, various terms associated with it have emerged to categorize and identify limitations in approximate computational techniques [38]. Although the correlation energy is a small fraction of the total energy of the system, typically of the order of 1%, it can be of the order of chemically important energies such as binding energies and electron affinities.

Diverse terminology associated with electronic correlation energies has been developed to elucidate the influence of electron motion within atoms, molecules, and polyatomic systems. In atoms, electron correlation can be classified into two primary forms: radial and angular [39, 40]. Radial correlation compels electrons to remain at a distance from nuclei, this kind of correlation is managed using basis functions with varying exponents. Angular correlation pertains to the likelihood of an electron’s presence on one side of a nucleus when another is on the opposite side, this is often described using high-angular-momentum functions.

Within diatomic molecules, the internuclear axis (z) defines a natural orientation of the coordinate system, and it turns to be appropriate to describe the correlation in terms of the z coordinates. In this case, the term axial (or "left-right") correlation has been introduced. It denotes the increased likelihood of an electron being in proximity to nucleus A when another electron is close to nucleus B. Alternatively, it is sometimes referred as the propensity of an electron to be located far from the z-axis due to the presence of another electron [41, 42]. The angular (or "equatorial") correlation in diatomics can be perceived as the complementary effect, aligning with the intuitive notion that, at large separations, an in-phase motion of the z-coordinates of the two particles and an out-of-phase motion of the x and y coordinates are energetically more favorable than their opposites [41]. For axial correlation description, electrons must be allowed to occupy the antibonding (σ_u) orbital. Meanwhile, for angular correlation, electrons must be allowed to occupy the higher energy π_u orbital. In both cases, a multiconfigurational wave function is required to adequately describe the system.

Finally, in polyatomic systems, electron correlation categorization is often based on criteria rooted in interelectronic distances, discerning between long-range (*lr*) and short-range (*sr*) interactions. The latter (*sr* correlation) arises from the singularities in the Hamiltonian due to the strong (instantaneous) interactions among electrons. This type of correlation is manifested in the appearance of electron-electron cusps in the exact wave function [17]. The behaviour of electrons near these cusps is too complicated to be described by small orbital expansions. Therefore, an acceptable description often requires the extension of the orbital basis, together with the inclusion of a large amount of excited configurations in the wave function. Following this strategy, the regions in the space where the wave function inaccurately represents the system can be reduced, but this often requires a large amount of computational effort. On the other hand *lr* interactions, are related to orbital degeneracies

or to van der Waals forces, and are manifested, for example, in the left-right and equatorial correlation emerging in the H_2 molecule [41, 42]. Fortunately, this type of correlation can be recovered, for instance, by wave functions with a relatively small number of determinants [17].

Another widely used classification of electron correlation is done in terms of dynamic and nondynamic correlation types [43, 44, 45, 46], this classification is often expressed within the framework of configuration interaction, where the overall electronic wave function is expanded in terms of different electronic configurations. Dynamic correlation (DC) is said to be included if the expansion contains a most dominant configuration and only small (but energetically important) contributions of other configurations. Similar to SR correlation, DC is attributed to the instantaneous electron repulsion, and is manifested from the inability of the HF wave function to model interelectronic cusps and dispersion interactions. Non-dynamic correlation (NDC), also termed static correlation, arises when there are two or more dominant configurations of similar importance, raising from near-degeneracy of electronic configurations (as represented by Slater determinants). DC is considered universal in the sense that instantaneous electron-electron repulsions are inherent to any system with at least two electrons, causing fluctuations in electron distribution. In contrast, NDC has been regarded as a system-specific quantity. The distinct effects of DC and NDC have been observed in terms of the electron density. The instantaneous repulsion of electrons that characterizes the DC, results in subtle localized changes in the electron density when compared to the HF description. NDC, however, induces significant, global alterations in the electron density due to the interactions among the various electronic configurations involved in its description. The impact of electron correlation on the electron density has been used to categorize the division of correlation energy into DC and NDC components [47, 48, 49].

As a final remark, an alternative definition of the correlation energy is described in the DFT formalism [50], where it is computed from a functional of the density $E_c^{DFT}[\rho]$. The exact DFT correlation energy can be computed by knowing the exact ground-state density and the exact correlation functional $E_c^{DFT}[\rho_{exact}]$. A relation of the correlation energy defined in Eq. 1.3 and that from DFT reads as $E_c \leq E_c^{DFT}$. However, since $E_c^{DFT} - E_c$ is small but non zero, self consistent DFT results for the correlation energy should not be compared directly with the conventional definition of Löwdin [51].

1.2.1 Methods for the treatment of electron correlation.

It is not possible to give a comprehensive account of all the literature regarded to electron correlation methods. However, according to the type of reference wave function used, two main classes of methods can be devised. The first is the single-reference-based (SR) approaches, which use a single Slater determinant as the reference wave function. The Møller-Plesser second-order perturbation theory (MP2) [52], CI and coupled cluster (CC) with single and double excitations (CISD or CCSD)[53, 54] are some typical examples. The second class is the multi-reference (MR) approaches, in which the reference wave function is composed of two or more determinants, examples are the multireference CI (MRCI) [55, 56, 57], the complete active space self-consistent field (CASSCF)[58] and its perturbative variant (CASPT2)[59].

Among WFT approaches, post-HF methods are designed to enhance the treatment of electron correlation beyond the limitations of the HF approach. These improvements are achieved through various approaches that use the HF wave function as the reference starting point.

Among post-HF strategies, in variational approaches, such as CI, the energy is always an upper bound to the true energy of the system. Particularly, the CI wave function is expressed as a linear combination of Slater determinants, including single and multiple excitations, with respect to the HF electronic configuration, to account for electron correlation effects. As the expansion increases, the energy is systematically closer to the exact solution. However, as mentioned at the beginning of this chapter, a full expansion is in most cases computationally intractable. Further details about CI formalism will be described in section 2.2.

Part of the difficulties related with the large computational effort encountered in WFT methods have been addressed in recent years. The improvement of computational power and the use of local-correlation, development of linear-scaling techniques [60, 61, 62, 63], and the recent developments of the density matrix renormalization group (DMRG) method [64] in multireference quantum chemistry offers the ability to describe static correlation in a large active space. All of this is bringing back the interest on WFT methods for medium- to large-size systems.

The reference wave function can introduce additional challenges. For instance, (single-reference) CI, MP[52], and CC [65, 66] approaches focuses on finding the total electronic energy of one electronic state based on the premise that the single-reference HF wave function is the dominant configuration. Unfortunately, the orbitals generated self-consistently in a mean field might be of poor relevance for the description of a multiconfigurational character system. The same problem arises in other nonvariational approximated methods, such as CC with singles and doubles (CCSD)[67], and CCSD with triples perturbative excitations (CCSD(T))[37], which can be applied to numerous molecular systems, providing accurate energies, but may fail in describing molecular dissociations.

In DFT, the problem turns to extract the required information from the electron density (a one-body quantity). Although formally DFT is exact, in practice the available approximations have several limitations. Most modern functionals can still have errors of 100 kcal/mol in relatively simple systems. These can be connected to systematic errors such as self-interaction error (SIE) and static correlation error in more complex systems [68]. SIE is considered one of the key failings of KS-DFT[69], it can be described as the result from the interaction of each electron with the entire electron density (including its own density). In KS-DFT, the energy is computed from a density functional approximation (DFA), so there is no way to precisely distinguish two-body Coulomb interactions from self-interaction, consequently, for a one-electron system, the SIE is equal to the entire Coulomb energy, this can be seen from the expression for the classical Coulomb integral

$$J[\rho] = \frac{1}{2}\int\int d\vec{\boldsymbol{r}}_1 d\vec{\boldsymbol{r}}_2 \frac{\rho(\vec{\boldsymbol{r}}_1)\rho(\vec{\boldsymbol{r}}_2)}{r_{12}} \tag{1.4}$$

where ρ is the electron density, $\vec{\boldsymbol{r}}$ the spatial coordinated of an electron, and r_{12} is the distance between two electrons. For a single electron, the Coulomb operator yields a non-zero value leading to a repulsive interaction with itself. This term also exists in WFT but is canceled out by the corresponding exchange term (at the HF level). Attempts to remove self-interaction from the Coulomb energy are performed by the exchange-correlation functional, as exchange is approximated in DFT, the imperfect cancellation leading to the SIE. In KS-DFT, SIE leads to artificially delocalized electron densities, resulting in an overestimation of the stability of molecules and ions. A widely used strategy to reduce SIE is through use of hybrid and range-separated hybrid functionals, since they include some fraction of exact HF exchange (which cancels with part of the Coulomb energy). However, the optimum

exchange-mixing fraction is system dependent [69].

A qualitative classification of the methods can be done according to their ability to describe DC and NDC. Regarding the treatment of DC, methods such as MP2, CCSD, CISD and DFAs are typical workhorses due to their performance. Due to the nature of the problem, NDC is accounted by multireference approaches such as CASSCF or DMRG. However, current efforts have been turned to the challenging description of systems presenting both DC and NDC. The simultaneous accurate treatment of both types of correlations is provided by very few methods such as CASPT2 or MRCI, and most recently in alternative approaches such as the adiabatic connection linked to MCSCF (AC-MCSCF) [70], the ΔNO method [71] and the global natural orbital functional (GNOF) approach [72].

Mingled WFT+DFT methods attempt to simultaneously address NDC and DC effects with manageable computational demands. This holds particular value for applications in modeling excited states, where challenges arise from the intrinsic nature of these states—almost all of which are open-shell states. In some cases, these states exhibit multiconfigurational character, especially those involving excitations to valence-like orbitals. Relevant challenges encompass convergence difficulties, points on the potential energy surface (PES) where near-degeneracies among different states occur. Additionally, a universal method for excited states should demonstrate accuracy across various types of excitations, including valence, Rydberg, charge-transfer, etc. [73].

Both SR and MR approaches can be independently integrated into mingled schemes. Given the vast amount of WFT approaches and DFAs available, literature presents numerous strategies for WFT+DFT combinations. Among the DFT mingled with single-configurational WFT approaches, the most representative involves combining density functional exchange, HF exchange, and DFT correlation, resulting in hybrid DFAs. When combining DFT with MC-WFT approaches, various choices can be made. For example, one can opt for CI calculations with Hamiltonian matrix elements obtained by using elements from DFT, utilize methods that incorporate density functional energy corrections to WFT energies, or combine DFT with CIS. A schematic representation of different DFT+MC-WFT choices and their features is summarized in Figure 1.1.

One of the major challenges for mingled WFT+DFT schemes, is related to the question, how to separate energy contributions covered by each of the two approaches?, i.e. how to avoid a partial double counting of electron correlation effects? [74, 75]. Strategies designed to avoid such problem characterize the mingled methods. For example, a particular class of DFT+MC-WFT, are those based on the range-separated formalism [22, 76, 77], in which short-range interactions are accounted by DFT and the long-range interactions by WFT [78, 79, 80]. This is done by splitting the two-electron interaction into short- and long-range terms. Under this strategy, the electron correlation double counting problem can be avoided (in principle), provided a correct separation of the short- and long-range interactions within the two different methodologies.

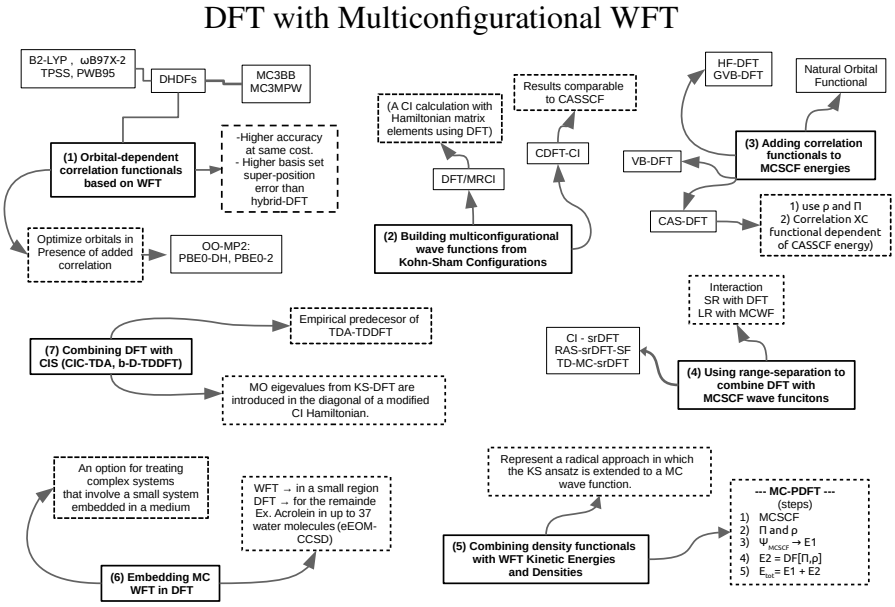


Figure 1.1: Some of the available DFT+MC-WFT mingled approaches. A very brief description of their features is included inside dashed-line squares.

Details on selected approaches able to simultaneously account for DC and NDC can be consulted in Apendix A of this thesis.

1.2.2 Desirable properties in an electronic structure method.

In the framework of electronic structure, method development stands as a dynamic and competitive research field, marked by diverse contributions from various groups (as noted in the previous section). Each proposed approach undergoes a comprehensive process, often spanning years to establish its validity. The process of development of a theoretical model chemistry can be divided into five stages [81, 1] (see Fig. 1.2). A successful candidate achieves certain general features, such as, ensuring validated accuracy (chemical accuracy, 1-2 kcal/mol, is globally accepted) for certain properties (e.g. heats of formation or ionization potentials). Its formulation must be as general as possible, avoiding specific procedures for particular molecules or particular symetries. In most cases, the computational aspect plays a significant role in determining the method's applicability to various chemical problems of interest. Thus, model's applicability should be achieved in a reasonable time and with reasonable computational resources.

In order to provide a focus for the subsequent content of this thesis, it is helpful to consider some properties that electronic structure methods should attempt to satisfy. In fact, the design of such a method faces challenges that go beyond achieving mere accuracy, since computational efficiency requires approximations that compromise transferability and accuracy. Over the years, various authors [1, 82, 83, 84] have established criteria that can be summarized in the following points

(1) The use of an empirical correction is undesirable.
(2) Independent on symmetry or specific choices of configurations.
(3) Size consistency, size extensivity and size intensivity.

(4) Invariant to classes of transformations, particularly unitary transformations among degenerate orbitals.
(5) Applicable to excited states and open shells.
(6) Should allow the generalization to include relativistic contributions.
(7) Inclusion of triple excitations.

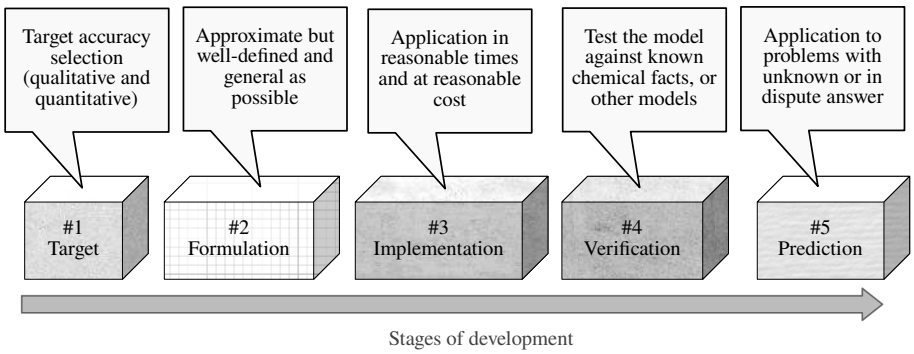


Figure 1.2: Pople's five stages of the development process for a theoretical model, adapted from Ref. [1].

Point (1) results from the performance observed from semiempirical approaches (methods that combine theoretical principles with empirical adjustments to parameters based on experimental data). These are limited by uncertainty over the consequences of the massive integral approximations and the large number of empirical parameters. Regarding point (2), methods designed for specific molecular cases can be made both fast and accurate. However, their transferability and accuracy for related molecules might be significantly compromised due to a lack of testing with reference data. Reference data could be incomplete, even when examined with a suitable measure of molecular and electronic similarity.

Three of the most vital considerations are the properties of size-consistency, size-extensivity, and size-intensivity (point 3). Size-consistency pertains to the ability to describe the PES of a given system. Originally, size-consistency was defined as the correct non-interacting description (in qualitative sense), what has been referred to as the "additive" separability of the wave function. If we imagine two H_2 molecules separated by a large distance, large enough that the molecules can be considered to be isolated one from the other, then the energy calculated for both molecules simultaneously should be exactly twice that from one isolated molecule of H_2. The size-consistency concept is more of a system-dependent problem than an inherent property of an electronic structure method. As for example, the restricted HF (RHF) solution for the dissociation of the H_2-H_2 dimer is size-consistent. But the RHF solutions are not size-consistent for closed-shell molecules dissociating into open-shell fragments (as in the dissociation of a H_2 molecule into two H atoms).

Size-consistency and size-extensivity are used interchangeably in the literature, however, the latter[82] is defined in a more mathematically formal way, referring to the method's correct (linear) scaling, i.e. an electronic structure method should provide size-extensive energies in the sense that the energy should grow linearly with the number of electrons in the system. The requirements for size-extensivity can be depicted as follows. Consider a system of two noninteracting fragments A and B, separated at an infinite distance for which the Hamiltonian operator can be written as

$$\hat{H}_{AB} = \hat{H}_A + \hat{H}_B, \tag{1.5}$$

where $\hat{H}_A$ is associated with the fragment A and $\hat{H}_B$ with fragment B. The exact solution for the combined system satisfies

$$\hat{H}_{AB} \left|AB\right\rangle = E_{AB} \left|AB\right\rangle, \tag{1.6}$$

while the exact solutions for the fragments are respectively

$$\hat{H}_A \left|A\right\rangle = E_A \left|A\right\rangle, \tag{1.7}$$

$$\hat{H}_B \left|B\right\rangle = E_B \left|B\right\rangle, \tag{1.8}$$

therefore, the exact compound wave function can be written as the product

$$\left|AB\right\rangle = \left|A\right\rangle \left|B\right\rangle, \tag{1.9}$$

and the exact energy for the compound system may be written as the sum,

$$E_{AB} = E_A + E_B. \tag{1.10}$$

These relations (1.9 and 1.10) constitute the requirements for size-extensivity [17]. In fact, both, size-consistency and size-extensivity are properties of the exact wave solution for the electronic Schrödinger equation. HF qualify as size-extensive, as well as MP2 and CC theories. On the other hand, in general, truncated CI methods with double or higher excitations are not size-extensive. Lack of this property implies that errors from the exact energy increase as more electrons enter the calculation. For systems composed of multiple non-interacting subsystems, size-extensivity ensures that the method generates the same energy, whether calculations are performed separately for each subsystem or simultaneously for the entire super-system.

Both, size-extensivity and size-intensivity were introduced in analogy with the properties in thermodynamics, where an extensive property scales with the size of the system and an intensive property is independent of the size of the system [85]. In the realm of theoretical chemistry, a method is said to be size-intensive if the excitation energy of a system, A, computed in isolation is equal to the excitation energy of A in the presence a non-interacting subsystem B,

$$E_{A*} - E_A = E_{A*\cup B} - E_{A\cup B}, \tag{1.11}$$

where labels $A*$ and A are the excited and ground states of subsystem A, respectively, as before, B label is for another subsystem that is in its electronic ground state and does not interact with A. A size-intensive method will show no change in the excitation energy as a function of the number of non-interacting molecules [86]. The failure in describing size-intensivity can be interpreted as the degradation of the accuracy of an individual excited state as additional fragments are present in the overall system.

The property of invariance under unitary transformations (4), meaning that the energy (or other properties) is unchanged by a rotation of the orbitals, has been subject of study by various authors [82, 87, 88]. For *ab-initio* methods, the molecular orbitals obtained as solutions to integro-differential one electron equations are often not unique. Different sets of orbitals can be related by different unitary transformations. Thus it is certainly desired that energy and other properties not be affected by the orbital representation. This susceptibility to change introduces certain degree of arbitrariness into the solutions. While the requirement of invariance was highlighted by Bartlett [82] in the framework of CC theory, it might be also of crucial importance for other theories. Kong [89] has demonstrated, through tensor analysis that the HF, single-reference CI, and CC methods are all invariant with respect to rotations within occupied and unocuppied orbital spaces. Similarly, the CASSCF and MRCI methods are invariant with respect to rotations in three spaces: core, active and inactive virtual orbital spaces [58] (a definition of such spaces can be consulted in section 2.3).

The orbital invariance problem is related with the calculation of excited states (5). This is because for open-shell cases (most of excited states) the invariance properties can be problematic, leading to complicate energy evaluation. However, the absence of invariance may not lead to significant issues, unless the orbital energies are essentially degenerated. In these cases, orbitals mix arbitrarily causing problems in schemes such as CASSCF followed by a CI calculation [87], or in variants of CC theory methods [82, 88].

Additional desirable properties (6 and 7) in an electronic structure method are related to the method's accuracy. Relativistic contributions are crucial for certain problems, as for example, the inclusion of spin-orbit coupling in the description of photophysical mechanisms [90], or kinetic and exchange corrections for heavy atoms [91].

In the calculations of energy differences such as heats of reaction, binding energies and activation energies to chemical accuracy, it is recognized (particularly in the realm of CC approaches) that, for high-accuracy results, the CCSD approach supplemented by a perturbative correction for connected triple substitutions is the most successful one in practical applications [92]. While part of the success of the (T) correction is due to a slight error compensation [93, 94], the accuracy delivered by CCSD(T), for much of computational chemistry, makes it the method of reference. Nevertheless, this level of accuracy is commonly associated with tremendous computational effort [61], making the inclusion of triple excitations not always desirable and sometimes not even necessary due to other sources of error [95].

CHAPTER 2

A review of theoretical background

Amongst the vast plethora of electronic structure methods, this chapter is dedicated to those approaches and analysis tools that are relevant to the development of this thesis. Further literature review on related relevant approaches is also included.

2.1 Hartree-Fock approximation

For a molecular system of N electrons and M nuclei described by the position vectors $\vec{\boldsymbol{r_i}}$ and $\vec{\boldsymbol{R_A}}$, respectively, the electronic Hamiltonian of Eq. 1.2 can be written in atomic units as

$$\hat{H}_{el} = -\sum_{i}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i}^{N}\sum_{A}^{M} \frac{Z_A}{r_{iA}} + \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} \tag{2.1}$$

where as indicated in the introduction, Z_A is the atomic number of the nucleus A, $1/r_{iA}$ the Coulomb attraction term between the ith-electron and the Ath-nucleus and $1/r_{ij}$ is the corresponding electron-electron interaction term. We can combine the one electron terms and rewrite this Hamiltonian 2.1 as

$$\hat{H}_{el} = \sum_{i}^{N} h(i) + \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} \tag{2.2}$$

The one-particle term ($H_0 = \sum_i^N h(i)$, in this case) can be exactly solved by a wave function with a structure that reduces to a simple separable product constructed of one-particle functions $\phi_i(\vec{\boldsymbol{r}}_i; \vec{\boldsymbol{R}}_A)$,

$$\Psi_0(r; R) = \prod_{i}^{N} \phi_i(\vec{\boldsymbol{r}}_i; \vec{\boldsymbol{R}}_A) \tag{2.3}$$

where the notation $\Psi_0(r; R)$ means a function of all particle coordinates indicated by r, that depend parametrically on the collective nuclear coordinates, R. By choosing each one-particle function as an eigenfunction such that $h(i)\phi_i(\vec{\boldsymbol{r}}_i) = \epsilon_i\phi_i(\vec{\boldsymbol{r}}_i)$, we would have

$$H_0\Psi_0(r;R) = E_0\Psi_0(r;R) \quad ; \quad E_0 = \sum_{i=1}^{N} \epsilon_i \tag{2.4}$$

We can choose a collection of one-particle functions to describe the individual behavior of electrons, one possible choice for such set of functions (a finite basis set) is the so-called atomic orbitals (AOs), describing electron probability distributions around an atomic nucleus, which is the usual choice within the quantum chemistry community, or alternatively, plane waves, which are typically used within the solid state community in periodic boundary conditions calculations. Several types of atomic orbitals can be used: Gaussian-type orbitals (GTOs), Slater-type orbitals (STOs), or numerical atomic orbitals (NAOs) [96]. For basis sets of comparable size NAOs are generally more accurate than GTOs, as demonstrated in various studies [97, 98, 99, 100]. GTOs are primarily employed in quantum chemistry packages because, in principle, STOs can be precisely expanded into GTOs [101], and because GTOs demonstrate computational efficiency in integral evaluation schemes[102, 103, 104, 105, 106].

GTOs will be the chosen basis set used throughout this thesis. The building blocks of a GTO are linear combinations of Gaussian functions that closely expand a given STO. The quality of this expansion is important since STOs represent analytical radial solutions to the hydrogenic Schrödinger equation. In fact, Slater orbitals are better from a theoretical point of view since they have the correct behavior near the nucleus (cusp) and the correct asymptotic long-range behavior. Consequently, a large number of Gaussians is needed for an accurate representation of the electronic system while typically fewer STOs than GTOs are needed to reach a certain level of accuracy [107]. Nonetheless, in practice, this drawback is counterbalanced by the quicker evaluation of molecular integrals involved in computations, permitting the use of a larger number of functions.

By including spin in each one-particle function (through multiplication by α or β spin functions), the corresponding functions are called *spinorbitals*

$$\chi_i(\vec{\boldsymbol{x}}_i) = \chi_i(\vec{\boldsymbol{r}}_i, \sigma_i) = \phi_i(\vec{\boldsymbol{r}}_i)\alpha(\sigma_i), \tag{2.5}$$

where σ is a coordinate in spin space, $\vec{\boldsymbol{x}}_i$ indicates both the space- and spin-coordinates ($\vec{\boldsymbol{x}}_i = \vec{\boldsymbol{r}}_i, \sigma_i$). When these form the basis set, the eigenfunction of the one-particle Hamiltonian (Eq. 2.3) is the so-called Hartree product (HP). HPs are eigenfunctions of any one-particle Hamiltonian. We can take advantage of this to solve Eq. 2.1, by expressing the two-particle term as the "average" of one-particle operators $u(\vec{\boldsymbol{r}}_i)$. Representing the average electron repulsion of all other electrons toward electron i, we can write 2.1 as

$$\begin{aligned} \hat{H}_{el} &= \sum_{i}^{N} [h(i) + u(\vec{\boldsymbol{r}}_i)] + \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} - \sum_{i}^{N} u(\vec{\boldsymbol{r}}_i) \\ &= \sum_{i}^{N} F(\vec{\boldsymbol{r}}_i) + V = \hat{H}_0 + V, \end{aligned} \tag{2.6}$$

with

$$V = \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} - \sum_{i}^{N} u(\vec{\boldsymbol{r}}_i) \tag{2.7}$$

$$\hat{F} = \sum_{i}^{N} [h(i) + u(\vec{\boldsymbol{r}}_i)], \tag{2.8}$$

where $\hat{F}(\vec{\boldsymbol{r}}_i)$ being the so-called Fock operator. For V small (clearly smaller than the two particle term), the Hamiltonian can be simplified to $\hat{H}_0 = \hat{F}$. This simplification reduces the N-particle problem to a set of one-particle eigenvalue problems,

$$\hat{H}_0 \chi_0(\vec{\boldsymbol{x}}_i) = \epsilon_i \chi_0(\vec{\boldsymbol{x}}_i). \tag{2.9}$$

The eigenvalues of this operator are known as the orbital energies, which are a measure of the energy that an electron (the sum of the kinetic energy and the average electrostatic interaction energy terms) in a given orbital.

The HP in the basis of spin orbitals is an uncorrelated wave function that does not account for the indistinguishability of electrons. Suitable solutions of the electronic Schrödinger equation must have the appropriate permutational symmetry, ($|\Psi_0|^2$ being invariant to the exchange of space and spin coordinates of any two electrons). To fulfill this requirement, the wave function is chosen to have the form of an antisymmetrized product, the Slater determinant,

$$\Psi_0 = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\vec{\boldsymbol{x}}_1) & \chi_2(\vec{\boldsymbol{x}}_1) & \dots & \chi_N(\vec{\boldsymbol{x}}_1) \\ \chi_1(\vec{\boldsymbol{x}}_2) & \chi_2(\vec{\boldsymbol{x}}_2) & \dots & \chi_N(\vec{\boldsymbol{x}}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\vec{\boldsymbol{x}}_N) & \chi_2(\vec{\boldsymbol{x}}_N) & \dots & \chi_N(\vec{\boldsymbol{x}}_N) \end{vmatrix} = |\chi_1(\vec{\boldsymbol{x}}_1)\chi_2(\vec{\boldsymbol{x}}_2)\dots\chi_N(\vec{\boldsymbol{x}}_N)\rangle\,, \tag{2.10}$$

being $1/\sqrt{N!}$ a normalization factor. Using a HP to form a Slater determinant introduces *exchange* effects, where the motion of two electrons with parallel spins is correlated, but for two electrons with opposite spins it remains uncorrelated. Additional eigenstates of $\hat{H}_0$ (denoted as Ψ_i^a) can be formed by substituting one occupied molecular orbital, $\chi_j(\vec{\boldsymbol{x}}_j)$, with an unoccupied orbital, $\chi_a(\vec{\boldsymbol{x}}_j)$ ($\Psi_i^a = |\chi_1(\vec{\boldsymbol{x}}_1)\dots\chi_a(\vec{\boldsymbol{x}}_j)\dots\chi_N(\vec{\boldsymbol{x}}_N)\rangle$). Various forms of Ψ_i^a, generated by different single orbital substitutions, represent potential 'configurations' derived from the ground state reference wave function. More complex configurations can be derived from multiple substitutions as well.

The independent particle model depicted above corresponds to the HF-SCF approximation when the Slater determinant becomes the "best" single determinantal approximation to Ψ, that is, the lowest-energy single determinant constructed from the variation of the orthogonal N-spin orbitals $\langle\chi_i|\chi_j\rangle = \delta_{ij}$. The energy expression defined as the expectation value of the Hamiltonian using the Slater determinant wave function corresponds to

$$\begin{aligned} E[\{\chi_i\}] = -\frac{1}{2}\sum_i \int d\vec{\boldsymbol{r}}\; \chi_i^*(\vec{\boldsymbol{r}}_i,\sigma_i)\nabla^2\chi_i(\vec{\boldsymbol{r}}_i,\sigma_i) + \int d\vec{\boldsymbol{r}}\; \rho(\vec{\boldsymbol{r}})v(\vec{\boldsymbol{r}}) \\ + \int d\vec{\boldsymbol{r}}\; d\vec{\boldsymbol{r}}'\; \frac{\rho(\vec{\boldsymbol{r}})\rho(\vec{\boldsymbol{r}}')}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} + E_x[\{\chi_i\}] \end{aligned} \tag{2.11}$$

the sum features the integration of spin variables (σ) contained in the spinorbitals, $v(\vec{\boldsymbol{r}})$ stands for the external potential formed by the nuclear charges, $\rho(\vec{\boldsymbol{r}})$ is the electron density given by

$$\rho(\vec{\boldsymbol{r}}) = \sum_j \sum_i |\chi_i(\vec{\boldsymbol{r}}, \sigma_j)|^2 \tag{2.12}$$

The exchange energy term is given by

$$E_x[\{\chi_i\}] = -\frac{1}{2} \int d\vec{\boldsymbol{r}}\, d\vec{\boldsymbol{r}}' \, \frac{\left|\gamma_\sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}')\right|^2}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|}, \tag{2.13}$$

with the one-particle density matrix, in this case given by

$$\gamma_\sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}') = \sum_i^{N_\sigma} \chi_i^*(\vec{\boldsymbol{r}}, \sigma)\chi_i(\vec{\boldsymbol{r}}'\sigma). \tag{2.14}$$

Hartree-Fock equations are a set of integro-differential equations for the effective one-particle orbitals, obtained from the variational theorem and minimizing the functional

$$L[\{\phi_i\}] = E[\{\phi_i\}] - \sum_{ij} \lambda_{ij} \left[\int d\vec{\boldsymbol{r}}\, \chi_i^*(\vec{\boldsymbol{r}}\sigma)\chi_j(\vec{\boldsymbol{r}}\sigma) - \delta_{ij} \right], \tag{2.15}$$

where orthogonality of the orbitals is ensured by the Lagrange multipliers λ_{ij}. Corresponding Euler-Lagrange equations are obtained from the variation of L with respect to the spin orbitals χ_i^*, χ_j and λ_{ij}, giving the corresponding Hartree-Fock equations,

$$\left[-\frac{1}{2}\nabla^2 + v(\vec{\boldsymbol{r}}) + \int d\vec{\boldsymbol{r}}' \, \frac{\rho(\vec{\boldsymbol{r}}')}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} \right] \chi_i(\vec{\boldsymbol{r}}\sigma) - \int d\vec{\boldsymbol{r}}' V_x^{HF}(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}')\chi_i(\vec{\boldsymbol{r}}'\sigma) = \epsilon_i \chi_i(\vec{\boldsymbol{r}}\sigma). \tag{2.16}$$

These equations unveil the main features of the HF approximation, which can be illustrated by replacing the integral kernel $V_x^{HF}(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}') = \gamma_\sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}')/\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|$, with a set of orbital dependent potentials

$$\left[-\frac{1}{2}\nabla^2 + v(\vec{\boldsymbol{r}}) + \int d\vec{\boldsymbol{r}}' \, \frac{\rho(\vec{\boldsymbol{r}}')}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} + v_i^{HF}(\vec{\boldsymbol{r}}\sigma) \right] \chi_i(\vec{\boldsymbol{r}}\sigma) = \epsilon_i \chi_i(\vec{\boldsymbol{r}}\sigma), \tag{2.17}$$

where

$$v_i^{HF}(\vec{\boldsymbol{r}}\sigma) = -\int d\vec{\boldsymbol{r}}' \frac{n_i(\vec{\boldsymbol{r}}\sigma, \vec{\boldsymbol{r}}'\sigma)}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|}, \tag{2.18}$$

and

$$n_i(\vec{\boldsymbol{r}}\sigma, \vec{\boldsymbol{r}}'\sigma) = \frac{\chi_i^*(\vec{\boldsymbol{r}}\sigma)\gamma_\sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}')\chi_i(\vec{\boldsymbol{r}}'\sigma)}{|\chi_i(\vec{\boldsymbol{r}}\sigma)|^2}. \tag{2.19}$$

For an occupied orbital, the term $\phi_i(\vec{\boldsymbol{r}}\sigma)\phi_i^*(\vec{\boldsymbol{r}}'\sigma)$ is dominant in $\gamma_\sigma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}')$, thus $n_i(\vec{\boldsymbol{r}}\sigma, \vec{\boldsymbol{r}}'\sigma) \approx |\chi_i(\vec{\boldsymbol{r}}\sigma)|^2$. Therefore, the term n_i for occupied orbitals makes

$$v_i^{HF}(\vec{\boldsymbol{r}}\sigma) \approx -\int d\vec{\boldsymbol{r}}' \, \frac{\left|\chi_i(\vec{\boldsymbol{r}}'\sigma)\right|^2}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|^2} \tag{2.20}$$

i.e., it cancels the density of spin orbital χ_i. In other words, each effective electron moves in an average field determined by all the other electrons, excluding the field of its own charge cloud. Additionally, when considering neutral atoms with nuclear charges Z, the asymptotic form of the total effective HF potential is different for occupied and virtual orbitals. In the former case, it resembles the following expansion [108]

$$v_{Total,i}^{HF} \sim -\frac{Z}{r} + \int d\vec{\boldsymbol{r}}' \, \frac{\rho(\vec{\boldsymbol{r}}') - \left|\chi_i(\vec{\boldsymbol{r}}'\sigma)\right|^2}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} \sim -\frac{Z - N + 1}{r} \quad (r \to \infty) \tag{2.21}$$

For unoccupied orbitals, the following expansion follows

$$v_{Total,i}^{HF} \sim -\frac{Z}{r} + \int d\vec{\boldsymbol{r}}' \, \frac{\rho(\vec{\boldsymbol{r}}')}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} \sim -\frac{Z - N}{r} \quad (r \to \infty) \tag{2.22}$$

These results show that for occupied orbitals, the asymptotic behavior of $v_{Total,i}^{HF}$ is $1/r$, but this is not the case for virtual orbitals, where the decay is more pronounced. Thus, under the HF approximation, the virtual spectrum of a neutral atom does not match the Rydberg-like series characteristic of particles bound by a Coulomb potential. This outcome suggests that the HF virtual orbitals will have a significant degree of diffusion, result that can be of importance when considering the utilization of HF orbitals in more sophisticated electron correlation approaches, such as those found in post-HF methods.

Finally, two crucial points need emphasis. First, for optimizing the utilization of spin symmetry in evaluating Hamiltonian matrix elements, it is beneficial to use the one-electron basis generated in spin-restricted SCF calculations. That is, the molecular orbitals are equal for α and β electrons, it enables the construction of spin-adapted linear combinations of Slater determinants, known as configuration state functions (CSFs). These CSFs are eigenfunctions of the S^2 operator, representing true electron spin multiplets. Second, concerning index notation in the rest of this chapter, we adhere to the standard convention: indices $p, q, r, s, \ldots$ denote arbitrary orbitals, $i, j, k, l, \ldots$ denote occupied orbitals (or inactive in the case of correlated methods) in the reference determinant (Ψ_0), t, u, v, w, x, y denote active orbitals in correlated methods, and $a, b, c, d, \ldots$ denote orbitals from the virtual space.

2.2 Configuration interaction formalism

Configuration interaction method is a straightforward application of the Ritz method of linear variations [53], yielding highly accurate wave functions for small and medium-sized molecular systems (for both open- and closed-shell cases). Under CI formalism, an exact solution to the many-electron problem is achieved by expanding the wave function within the complete set of determinants generated through the distribution of all electrons across all orbitals (*full* CI *expansion*). By introducing constraints to the construction of configuration expansions, the CI wave function can be truncated, resulting in various approximations approaching the Full Configuration Interaction (FCI) limit.

An arbitrary normalizable N-electron wave function $\Psi = \Psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_N)$ taken at fixed values of space and spin coordinates of the first $N-1$ electrons ($\vec{\boldsymbol{x}}_q = \vec{\boldsymbol{r}}_q, \sigma_q = \text{const}$), can be expanded in a complete set of one-electron functions, $\chi_p = \chi_p(\vec{\boldsymbol{x}}), p = 1, 2, \ldots, \infty$ as

$$\Psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_N) = \sum_{i_N=1}^{\infty} c_{i_N} \chi_{i_N}(\vec{\boldsymbol{x}}_N) \tag{2.23}$$

The coefficients of this expansion are functions of the coordinates of the other $N-1$ electrons ($p = 1, 2, \ldots, N-1$). As such, they should also be normalizable functions of their arguments. Therefore, we can repeat the expansion, and for fixed values of $\vec{\boldsymbol{x}}_p$, $p = 1, 2, \ldots, N-2$, we may expand $c_{p_N} \equiv c_{p_N}(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_{N-1})$ into series as a function of $\vec{\boldsymbol{x}}_{N-1}$; we get

$$\Psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_N) = \sum_{p_{N-1}, p_N=1}^{\infty} c_{p_{N-1}p_N}(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_{N-2}) \chi_{p_{N-1}}(\vec{\boldsymbol{x}}_{N-1}) \chi_{p_N}(\vec{\boldsymbol{x}}_N). \tag{2.24}$$

Continuous repetition of this procedure will allow us to expand the total wave function as

$$\Psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_N) = \sum_{p_1, p_2, \ldots, p_N=1}^{\infty} c_{p_1 p_2 \ldots p_N} \chi_{p_1}(\vec{\boldsymbol{x}}_1) \chi_{p_2}(\vec{\boldsymbol{x}}_2) \ldots \chi_{p_N}(\vec{\boldsymbol{x}}_N), \tag{2.25}$$

where coefficients $c_{p_1 p_2 \ldots p_N}$ become independent of the electronic coordinates. The antisymmetry of the wave function leads to some restrictions on the values of the coefficients in the expansion. For instance, with respect to the interchange of electrons "1" and "2" it follows that

$$c_{p_1 p_2 \ldots p_N} = -c_{p_2 p_1 \ldots p_N} \tag{2.26}$$

and so forth. Slater determinants, which can be built from a given set of spin orbitals, are constructed by collecting those terms that contain the same spin orbitals, and differ only in the order of the electron coordinates and in the order of the subscripts $p_1, p_2, \ldots, p_N$ of the coefficients $c_{p_1 p_2 \ldots p_N}$. Requesting the subscripts p_q to form an increasing sequence ($p_1 < p_2 < \cdots < p_N$), the set of N indices p_q is called the Ith "ordered configuration". Within this structure, the *CI* wave function consists of a linear combination of the set of Slater determinants Ψ_I, corresponding to the different ordered configurations [109],

$$\Psi = \sum_I C_I \Psi_I. \tag{2.27}$$

If the one-electron functions form a complete set in the one-electron Hilbert space, then the determinants corresponding to all the ordered configurations form a complete set in the linear space of the antisymmetric N-electron functions. Additionally, if the one-electron functions are orthonormalizd, then the determinants built up from them are orthonormalized as well. In practice, as we start from a finite number of spin orbitals, χ_p, the determinants corresponding to the ordered configurations form a complete basis in the subspace of antisymmetric wave functions.

The CI method consists in the variational determination of the expansion coefficients appearing in Eq. 2.27,

$$\frac{\partial}{\partial C_I}\frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = 0, \tag{2.28}$$

This requires the solution of a secular equation corresponding to a linear variational problem, equivalent to an eigenvalue problem for the coefficients and the energy,

$$\mathbf{HC} = E\mathbf{C}, \tag{2.29}$$

where **C** is a vector containing the expansion coefficients. The CI expansion is written out in terms of one-electron spin orbitals, so it does not contain explicitly the interelectronic distances. Consequently, the behavior of the wave function in the cases when two electrons are close to each other is described only as a combined effect of many different terms in the CI expansion. This is the underlying reason why the CI expansion converges rather slowly, specially as far as the local properties of the wave function are concerned. This lack in providing a compact description of electronic correlation is the principal limitation of the CI method. Consequently, a considerable amount of computational effort is needed to perform accurate CI calculations for large molecules.

Standard CI implementations start by forming several Slater determinants all having the required spatial and spin symmetries, that differ from the dominant determinant of the ground state through promotion of one, two, three, or more electrons. Once the CI code is instructed as to which *single*, *double*, etc. promotions to include, it forms the relevant Slater determinants an then forms the Hamiltonian matrix elements between all pairs of determinants. These matrix elements are all expressed in terms of the MO-based integrals. The dimension of the CI Hamiltonian matrix thus formed is $N_{det} \cdot N_{det}$ and to find N_{states} eigenvalues requires $N_{states} \cdot N_{det}^2$ steps, which can be out of control if N_{det} becomes very large. For this reason, some CI-based calculations involve careful selection of which determinants to include in the procedure [14], and, in practice, most CI methods are restricted to finding only the lowest eigenvalue of $\hat{H}_{el}$ or few of them. This is key to make the method feasible, as FCI methodologies rarely give all eigentstates of $\hat{H}_{el}$.

This can be followed by a further step in which the one-electron (and perhaps two-electron) reduced density matrices are computed from the eigenvectors. These matrices can be stored and used to obtain expectation values for the corresponding electronic properties. Reduced transition density matrices, involving two different eigenvectors, can also be obtained and used for the calculation of transition moments between two electronic states [53]. A vastly mentioned solution to the problem of computational cost relies on the systematic truncation of the CI expansion, giving the so-called CIS, CISD, CISDT, ..., etc. approaches. However, truncation introduces a lack of size-extensivity in the wave function (which can be repaired by Davidson's quadratic renormalization [110, 111]). The CI ansatz is also taken by other methods, by restricting the FCI solution to a portion of the orbital space (the so-called active space). Computational effort can be diminished giving rise to application for larger systems, the following sections are devoted to describe single reference methods of this type.

2.2.1 CIPSI algorithm

Selected configuration interaction approaches rely on the same principle as the usual CI approaches, except that determinants are not chosen *a priori* based on an occupation or excitation criterion, but selected on the fly among the entire set of determinants. This selection is based on their estimated contribution to the FCI wave function.

A selective procedure is important from the fact that even inside a predefined subspace of determinants, only a small number significantly contributes to the wave function. Among other methodologies, for comparative purposes and testing of the methods developed in this thesis, we recur to the usage of CIPSI as introduced by Huron and co-workers [112] (and implemented in Quantum package [113]).

CIPSI is an iterative CI selection algorithm that allows us to perturbatively select determinants from the FCI space. CIPSI wave function is built iteratively, starting from a reference wave function composed of a linear combination of Slater determinants, Ψ_I, spanning a internal space, S.

$$\Psi^{CIPSI} = \sum_{\Psi_I \in S} C_I \Psi_I. \tag{2.30}$$

Typically, this reference is the lowest-energy HF single determinant or a CASSCF wave function. External determinants ($\Psi_\alpha \notin S$) are then computed and added to the expansion based on their effective second-order perturbative energy contribution [114].

At every iteration, the Hamiltonian is diagonalized in the reference space S to obtain the variationally minimized energy E_{ref} and expansion coefficients C_I. Then the individual contributions to the energy from all the members in the set $\{\Psi_\alpha\}$ are estimated from perturbation theory as

$$e_\alpha^{(2)} = \frac{\left\langle \Psi^{CIPSI} \middle| \hat{H} \middle| \Psi_\alpha \right\rangle^2}{E_{ref} - \langle \Psi_\alpha | \hat{H} | \Psi_\alpha \rangle}. \tag{2.31}$$

If this second-order energy is greater than a fixed or iteratively modifiable threshold, Ψ_α is selected for the next iteration. The sum of all these contributions is

$$E_{PT2} = \sum_\alpha e_\alpha^{(2)}, \tag{2.32}$$

which is an estimate of the total missing correlation energy. The CIPSI energy of the current iteration is therefore,

$$E_{CIPSI} = E_{PT2} + E_{ref}, \tag{2.33}$$

the subset of all $\{\Psi_\alpha^*\}$ determinants with the largest contributions is then added to the variational space. Selection iterations are repeated up to a given number of determinants (typical calulations can include millions) until some other criterion is met (e.g. reasonably convergence of the CIPSI energy). Within this algorithm, the quality of the wave function is systematically improved and converges toward the FCI solution.

2.3 Complete active space approaches

When designing configuration spaces smaller than FCI, a critical consideration is the treatment of both static and DC (DC) energy. Addressing static correlation involves retaining, within the truncated expansion, dominant configurations from the FCI expansion, along with nearly degenerate configurations. On the other hand, DC energy treatment involves introducing many configurations generated by excitations beyond the reference space—such as single, double, and triple excitations.

Configurations chosen for an accurate representation of static correlation are often termed 'reference configurations' of the CI wave function, collectively forming the 'reference space.' In CAS single reference methods, this reference space typically consists of a single configuration (usually a HF determinant denoted as Ψ_0). The system's orbital space is divided into 'active' orbitals with variable electron occupation, 'inactive' orbitals that are doubly occupied, and unoccupied orbitals composing the 'virtual' space. While inactive electrons are treated in a mean-field manner, exact solutions to the FCI problem are obtained within the smaller active space, procedure that is referred as CASCI method.

Further improvements over the CASCI wave function can be achieved by imposing certain restrictions to the electrons in the orbital spaces. There is a whole family of methods for different restrictions (e.g. RASCI [115], RASSCF [116], GASSCF [117], LASSCF [118]). For instance, in the case of the restricted active space (RAS) wave functions, active orbitals are classified into subcategories (RAS1, RAS2, and RAS3). In constructing the RAS wave function, constraints are set on the allowed number of electrons in RAS1 and RAS3, while no restrictions apply to the occupations in RAS2. As an example, a multi-reference singles and doubles CI (MRSDCI) wave function with a CAS reference space can be realized within the RAS framework, where RAS1 corresponds to inactive orbitals, RAS2 to active CAS orbitals, and RAS3 to secondary orbitals. RAS1 can have at most two holes, and RAS3 can have at most two electrons, with no constraints on RAS2, except that the total number of electrons remains equal to N. This unified approach to constructing configuration expansions simplifies wave function design and optimization significantly. In fact, the use of CAS or RAS methods allows moderately sized systems to be studied as long as the active space is relatively small. Without considering symmetry reduction of computational cost, the total number of possible Slater determinants for an active space with M spatial orbitals, N_α up spin electrons and N_β down spin electrons is

$$N_{Total} = \frac{M!}{N_\alpha!(M-N_\alpha)!}\frac{M!}{N_\beta!(M-N_\beta)!}. \tag{2.34}$$

This factorial growth establishes a bottleneck in the implementation and use of CAS-based approaches. Modern computing architectures can handle active spaces of approximately 18 electrons in 18 orbitals ($\sim 2 \times 10^9$ determinants) [119]. This limit can be extended with the use of massively parallel supercomputer implementations, partially up to 22 electrons in 22 orbitals ($\sim 5 \times 10^1 1$ determinants) [120]. More recently, adaptive sampling methodologies have enabled the exploration of extremely large active spaces, performing calculations with ~ 50 active electrons in ~ 50 active orbitals through the Adaptive Sampling Configuration Interaction (ASCI) method, strategy employed to study large molecules such as polyradical periacenes and iron porphyin (see reference [121]).

Regarding the orbitals in a CI expansion, several crucial points should be noted. For effectively treating static correlation, orbital determination schemes and reference configurations

is essential [122], as multiple configurations may contribute similarly to the final state. It is necessary to mantain a balanced treatment of all significant reference configurations. Along a PES, the relative importance of various configurations may change, necessitating the inclusion of the full set of reference configurations in orbital optimization to prevent bias in the description of the PES. Moreover, to recover a substantial portion of DC energy through CI requires a large number of excited configurations, making simultaneous optimization of orbitals and configuration coefficients challenging, if not infeasible. In the upcoming sections, we will dive into details about CASCI, RASCI, and CASSCF methods.

2.3.1 Complete active space configuration interaction (CASCI)

The CASCI wave function is constructed as a linear combination of configurations resulting from all the possible electronic distributions within the active space and the expansion coefficients are solved in a variational procedure. In this case, no restrictions are placed on the occupations of the active orbitals, but the remaining orbitals are either always doubly occupied (inactive set), or always unoccupied (secondary set). The orbitals remain fixed without further optimiztion. For a CASCI expansion the energy is invariant to rotations within the inactive, active, and virtual orbital subspaces.

Upon defining the active space, the CASCI Hamiltonian is formed and usually expressed within the basis of active space orbitals. An initial Fock matrix is then estimated, often based on these active space orbitals. Subsequently, diagonalization of this Fock matrix takes place within the active space, yielding molecular orbitals and their respective orbital energies. This diagonalization process ensures that the Fock matrix within the active space either becomes diagonal or approaches diagonalization. During this step, the inactive space orbitals typically remain unchanged.

In order to show part of the formalism. We express Hamiltonian 2.1 in second quantization (spin-orbital representation) as

$$\begin{aligned}\hat{H} = &-\frac{1}{2}\sum_{\sigma}\int d\vec{\boldsymbol{r}}\,\hat{\psi}^{\dagger}_{\sigma}(\vec{\boldsymbol{r}})\nabla^2\hat{\psi}_{\sigma}(\vec{\boldsymbol{r}}) \\ &+\sum_{\sigma}\int d\vec{\boldsymbol{r}}\,\hat{\psi}^{\dagger}_{\sigma}(\vec{\boldsymbol{r}})v(\vec{\boldsymbol{r}})\hat{\psi}_{\sigma}(\vec{\boldsymbol{r}}) \\ &+\frac{1}{2}\sum_{\sigma}\sum_{\sigma'}\int d\vec{\boldsymbol{r}}\,d\vec{\boldsymbol{r}}'\,\hat{\psi}^{\dagger}_{\sigma}(\vec{\boldsymbol{r}})\hat{\psi}^{\dagger}_{\sigma'}(\vec{\boldsymbol{r}}')w(\vec{\boldsymbol{r}},\vec{\boldsymbol{r}}')\hat{\psi}_{\sigma'}(\vec{\boldsymbol{r}}')\hat{\psi}_{\sigma}(\vec{\boldsymbol{r}})\end{aligned} \tag{2.35}$$

where σ are the spin coordinates of the electrons, $v(\vec{\boldsymbol{r}})$ the external potential created by nuclei and $w(\vec{\boldsymbol{r}},\vec{\boldsymbol{r}}')$ the potential among electrons (for Coulombic systems $w = 1/|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'|$). As previously indicated, $\vec{\boldsymbol{x}}_i = (\vec{\boldsymbol{r}}_i, \sigma_i)$. The action of the ladder operators $\hat{\psi}_{\sigma}(\vec{\boldsymbol{r}})^{\dagger}$ and $\hat{\psi}_{\sigma}(\vec{\boldsymbol{r}})$ on an N-electron state Ψ_N is defined as

$$\hat{\psi}^{\dagger}_{\sigma}(\vec{\boldsymbol{r}})\Psi_{N+1}(\vec{\boldsymbol{x}}_1,\ldots,\vec{\boldsymbol{x}}_N) = \frac{1}{\sqrt{N+1}}\hat{A}[\delta(\vec{\boldsymbol{x}} - \vec{\boldsymbol{x}}_{N+1})]\Psi_N(\vec{\boldsymbol{x}}_1,\ldots,\vec{\boldsymbol{x}}_N), \tag{2.36}$$

where $\delta(\vec{\boldsymbol{x}} - \vec{\boldsymbol{x}}_i) = \delta(\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}_i)\delta_{\sigma\sigma_i}$. The operator $\hat{\psi}^{\dagger}_{\sigma}(\vec{\boldsymbol{r}})$ adds a particle with spin σ at position $\vec{\boldsymbol{r}}$ to the N-particle state, while $\hat{\psi}_{\sigma}(\vec{\boldsymbol{r}})$ anhilates a particle with spin σ from the N-particle state. Operator $\hat{A}$ is the antisymmetrization operator defined as

$$\hat{A}\Psi(x_1,\ldots,x_N) = \sum_{P\in S_N} \mathrm{sign}(P)\Psi(x_{P(1)},\ldots,x_{P(N)}) \tag{2.37}$$

where the summation runs over all permutations P of the group of permutations S_N of N elements. The expectation value of the electronic Hamiltonian ($\hat{H}$) with state $|\Psi\rangle$ can be written as

$$\begin{aligned}\langle\Psi|\hat{H}|\Psi\rangle = &-\frac{1}{2}\sum_{\sigma}\int d\vec{\boldsymbol{r}}\,\nabla^2_{r'}\gamma_\sigma(\vec{\boldsymbol{r}},\vec{\boldsymbol{r}}')|_{r'=r}\\ &+\sum_{\sigma}\int d\vec{\boldsymbol{r}}\,\rho_\sigma(\vec{\boldsymbol{r}})v(\vec{\boldsymbol{r}})\\ &+\frac{1}{2}\sum_{\sigma}\sum_{\sigma'}\int d\vec{\boldsymbol{r}}\,d\vec{\boldsymbol{r}}'\,\Gamma_{\sigma\sigma'}(\vec{\boldsymbol{r}},\vec{\boldsymbol{r}}')w(\vec{\boldsymbol{r}},\vec{\boldsymbol{r}}').\end{aligned} \tag{2.38}$$

The electronic energy in Eq. 2.38 is expressed in terms of the electron density ρ_σ, the one-particle reduced density matrix γ_σ (1-RDM) and the diagonal two-particle reduced density matrix $\Gamma_{\sigma\sigma'}$ (2-RDM) defined in our notation as

$$\rho(\vec{\boldsymbol{r}}) = \langle\Psi|\sum_{\sigma}\hat{\psi}^\dagger_\sigma(\vec{\boldsymbol{r}})\hat{\psi}_\sigma(\vec{\boldsymbol{r}})|\Psi\rangle = N\sum_{\sigma}\int d\vec{\boldsymbol{x}}_2\ldots d\vec{\boldsymbol{x}}_N\left|\Psi(\vec{\boldsymbol{r}}\sigma,\ldots,\vec{\boldsymbol{x}}_N)\right|^2 \tag{2.39}$$

$$\begin{aligned}\gamma_\sigma(\vec{\boldsymbol{r}},\vec{\boldsymbol{r}}') =& \langle\Psi|\hat{\psi}^\dagger_\sigma(\vec{\boldsymbol{r}}')\hat{\psi}_\sigma(\vec{\boldsymbol{r}})|\Psi\rangle\\ =&N\int d\vec{\boldsymbol{x}}_2\ldots d\vec{\boldsymbol{x}}_N\Psi^*(\vec{\boldsymbol{r}}'\sigma,\vec{\boldsymbol{x}}_2,\ldots,\vec{\boldsymbol{x}}_N,)\Psi(\vec{\boldsymbol{r}}\sigma,\vec{\boldsymbol{x}}_2,\ldots,\vec{\boldsymbol{x}}_N)\end{aligned} \tag{2.40}$$

$$\begin{aligned}\Gamma_{\sigma\sigma'}(\vec{\boldsymbol{r}},\vec{\boldsymbol{r}}') =& \langle\Psi|\hat{\psi}^\dagger_\sigma(\vec{\boldsymbol{r}})\hat{\psi}^\dagger_{\sigma'}(\vec{\boldsymbol{r}}')\hat{\psi}_{\sigma'}(\vec{\boldsymbol{r}}')\hat{\psi}_\sigma(\vec{\boldsymbol{r}})|\Psi\rangle\\ =&N(N-1)\int d\vec{\boldsymbol{x}}_3\ldots d\vec{\boldsymbol{x}}_N\left|\Psi(\vec{\boldsymbol{r}}\sigma,\vec{\boldsymbol{r}}'\sigma'\vec{\boldsymbol{x}}_3\ldots\vec{\boldsymbol{x}}_N)\right|^2,\end{aligned} \tag{2.41}$$

for an overview of the properties of density matrices see references [123, 124]. Here, it is worth noting a widely used quantity: the on-top pair density defined for state I by

$$\Pi_I(\vec{\boldsymbol{r}}) = N(N-1)\int|\Psi_I(\vec{\boldsymbol{r}}_1,\vec{\boldsymbol{r}}_2,\vec{\boldsymbol{r}}_3,\ldots,\vec{\boldsymbol{r}}_N)|^2 d\vec{\boldsymbol{r}}_1 d\vec{\boldsymbol{r}}_2\ldots d\vec{\boldsymbol{r}}_N|_{\vec{\boldsymbol{r}}_1=\vec{\boldsymbol{r}}_2=\vec{\boldsymbol{r}}} \tag{2.42}$$

which corresponds to the probability density of finding two electrons at a point δ in space. As will be shown, $\Pi_I(\delta)$ is used in various works attempting to introduce the simultaneous treatment of both DC and NDC. Equivalently, in the orbital representation, the 1-RDM and 2-RDM are described by the ${}^nD$ tensor named n-density matrix (n-DM), a $2n$-index tensor of dimension K^n (K being the size of the orbital basis set), the orbital representations for the spinless 1-RDM and 2-RDM are

$$\gamma_\sigma(\vec{\boldsymbol{r}}'_1;\vec{\boldsymbol{r}}_1) = \sum_{pq}^{K}{}^1D_{pq}\phi^*_p(\vec{\boldsymbol{r}}'_1)\phi_q(\vec{\boldsymbol{r}}_1) \tag{2.43}$$

$$\gamma_\sigma(\vec{\boldsymbol{r}}_1', \vec{\boldsymbol{r}}_2'; \vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2) = \sum_{pqrs}^{K} {}^2D_{pqrs}\phi_p^*(\vec{\boldsymbol{r}}_1')\phi_q^*(\vec{\boldsymbol{r}}_2')\phi_r(\vec{\boldsymbol{r}}_1)\phi_s(\vec{\boldsymbol{r}}_2). \tag{2.44}$$

In this representation, the electronic energy (Eq. 2.39) can be written in terms of the 1-RDM and 2-RDM as

$$E = \sum_{pq} {}^1D_{pq}h_{pq} + \frac{1}{2}\sum_{pqrs} {}^2D_{pqrs}g_{pqrs}, \tag{2.45}$$

where the p, q, r, s indices are labels for different orbitals, ${}^1D_{pq}$ and ${}^2D_{pqrs}$ the corresponding one- and two-electron probability amplitudes, and h_{pq} and g_{pqrs} the one- and two-electron integrals defined as

$$h_{pq} = \int d\vec{\boldsymbol{x}}\phi_p^*(\vec{\boldsymbol{x}})\left(-\frac{1}{2}\nabla^2 - \sum_A \frac{Z_A}{r_A}\right)\phi_q(\vec{\boldsymbol{x}}), \tag{2.46}$$

$$g_{pqrs} = \int\int d\vec{\boldsymbol{x}}_1\, d\vec{\boldsymbol{x}}_2 \frac{\phi_p^*(\vec{\boldsymbol{x}}_1)\phi_r^*(\vec{\boldsymbol{x}}_2)\phi_q(\vec{\boldsymbol{x}}_1)\phi_s(\vec{\boldsymbol{x}}_2)}{r_{12}}, \tag{2.47}$$

where r_A stands for the electron-nuclear separation. The generalized Fock matrix is defined as

$$F_{mn} = \sum_q {}^1D_{mq}h_{nq} + \sum_{qrs} {}^2D_{mqrs}g_{nqrs}. \tag{2.48}$$

This is general and valid for any kind of wave function, however, because of the different weights on the one- and two-electron parts, the electronic energy cannot be evaluated as the trace of this matrix. To illustrate this, first, we can write Eq. 2.45 in the form

$$E = \frac{1}{2}\sum_{pq} {}^1D_{pq}h_{pq} + \frac{1}{2}\sum_{pl} {}^1D_{pl}h_{pl} + \frac{1}{2}\sum_{pqrs} {}^2D_{pqrs}g_{pqrs}, \tag{2.49}$$

where l is a dummy index introduced here for convenience. This equation can be further simplified by introducing a Kronecker delta function δ_{pq} and grouping corresponding indices

$$E = \frac{1}{2}\sum_{pq}\left({}^1D_{pq}h_{pq} + \delta_{pq}\left[\sum_l {}^1D_{pl}h_{ql} + \sum_{lrs} {}^2D_{plrs}g_{qlrs}\right]\right), \tag{2.50}$$

comparing this expression with the definition of the generalized Fock matrix, Eq. 2.48, leads to the formulation of the energy as

$$E = \frac{1}{2}\sum_{pq}\left({}^1D_{pq}h_{pq} + \delta_{pq}F_{pq}\right). \tag{2.51}$$

Instead, the energy evaluation can be simplified by taking into account the different occupancies of the inactive, active and virtual orbitals. We first note that, when assuming doubly occupied orbitals in a restricted formulation, the elements of the density matrices ${}^1D_{pq}$ and

${}^2D_{pqrs}$ vanish in the presence of virtual indices. When there is one inactive index, i, the evaluation turns to be

$$ {}^1D_{iq} = 2\delta_{iq} \tag{2.52}$$

$$ {}^2D_{pqis} = 2\delta_{is}\,{}^1D_{pq} - \delta_{iq}\,{}^1D_{ps}. \tag{2.53}$$

Thus, we can split the generalized Fock matrix into two components, one involving purely the inactive density and the other involving the J and K matrices built from the active space. When the first index corresponds to an inactive orbital and the second is not restricted

$$ F_{in} = 2\big({}^I F_{ni} + {}^A F_{ni}\big), \tag{2.54}$$

where the inactive and active Fock matrices are introduced respectively as

$$ {}^I F_{mn} = h_{mn} + \sum_i (2g_{mnii} - g_{miin}), \tag{2.55}$$

$$ {}^A F_{mn} = \sum_{vw} (g_{mnvw} - \frac{1}{2} g_{mwvn}). \tag{2.56}$$

Next, when the first index is active (and the second is general), we obtain

$$ F_{vn} = \sum_w^I F_{nw}\gamma_{vw} + Q_{vn}, \tag{2.57}$$

where the auxiliay Q matrix is given by

$$ Q_{vm} = \sum_{wxy} \Gamma_{vwxy} g_{mwxy}, \tag{2.58}$$

Finally, if the first index is virtual, $F_{an} = 0$. Inactive and active Fock matrices are symmetric. This formulation of the generalized Fock matrix is quite useful because it only requires density matrices with all indices active and two-electron integrals in the MO basis with three indices active and one general index, greatly reducing the storage and computational cost of the MO transformation.

2.3.2 Restricted active space configuration interaction (RASCI)

In a similar fashion to CASCI, the single reference RASCI family of methods emerges from splitting the orbital space of a reference configuration (Ψ_0) into three different subspaces: RAS1, RAS2 and RAS3. With RAS1 (RAS3) orbitals being doubly occupied (virtual) in Ψ_0. As in other single-reference CI methods, the RASCI wave function for a target state $\left|\Psi^{RASCI}\right\rangle$ can be obtained by the action of an excitation operator ($\hat{R}$) on a reference wave function as

$$ \left|\Psi^{RASCI}\right\rangle = \hat{R}\left|\Psi_0\right\rangle = \sum_I C_I \left|\Psi_I\right\rangle, \tag{2.59}$$

the excitation operator $\hat{R}$ can be expanded as

$$\hat{R} = \hat{r}_0 + \hat{r}_h + \hat{r}_p + \hat{r}_{hp} + \hat{r}_{2h} + \hat{r}_{2p} + ..., \tag{2.60}$$

with $\hat{r}_0$ containing all possible excitations within RAS2, (reduced FCI) and the rest of the terms in Eq. 2.60 generate configurations with increasing number of holes (h subindex) in RAS1 and/or particles (p subindex) in RAS3. By setting the dimensions of these two orbital spaces, RAS2 is automatically defined (see figure 2.1).

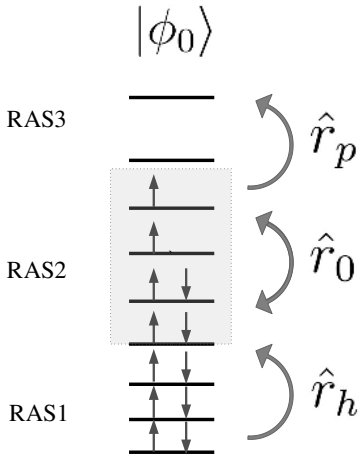


Figure 2.1: Representation of the three RAS spaces and the action of the $\hat{r_0}$, $\hat{r_h}$ (*hole*) and $\hat{r_p}$ (*particle*) terms of the excitation operator on the reference state ($|\Psi_0\rangle$), here exemplifying a triplet state with an active space of six electrons in four orbitals (RAS2(6e,4o)).

The expansion of $\hat{R}$ to all possible number of holes and particles is equivalent to FCI. The CI amplitudes (C_I) and the state energies are obtained by iterative diagonalization of the corresponding RASCI Hamiltonian:

$$\sum_J \left\langle \Psi_I \middle| \hat{H} \middle| \Psi_J \right\rangle C_J = E C_I. \tag{2.61}$$

Truncation of the excitation operator from Eq. 2.60 leads to restrictions on the number of holes in RAS1 and particles in RAS3, giving different levels of approximation. The simplest approach is to truncate $\hat{R}$ to its zeroth order term (reducing to the CASCI ansatz), while systematic improvements are achieved by enabling $\hat{r}_h$ and $\hat{r}_p$ terms in the $\hat{R}$ expansion. Other flavors of RASCI can be defined based on the nature of $\hat{R}$: excitation energy (EE), spin-flip (SF), performing excitations of spin-α electrons to empty spin-β orbitals (or vice-versa), and the ionization potential (IP) and electron attachment (EA) operators connecting Hilbert spaces with different number of electrons [125, 126]. In any case, spin completeness is ensured independently of the employed reference configuration and the nature of $\hat{R}$. The flexibility of the RASCI wave function naturally allows the treatment of highly correlated systems, in particular the RASCI(h, p) approach in combination with the SF excitation operator has shown to provide a good compromise between accuracy and computational cost [127].

2.3.3 Multiconfigurational Self-Consistent Field (MCSCF)

Since the 1930s, Multiconfigurational Self-Consistent Field (MCSCF) methods have evolved as a significant advancement in describing molecular systems where electron correlation plays a crucial role [128]. They are generalizations of the Hartree-Fock model built upon the idea of carry out a CI calculation where orbitals are variationally optimized simultaneously with the coefficients of the electronic configurations. In MCSCF theory, the wave function is constructed as a linear combination of Slater determinants or CSFs, of the form

$$|\mathbf{k},\,\mathbf{C}\rangle = \exp\left(-\hat{k}\right) \sum_{I} C_I \,|\Psi_I\rangle\,, \tag{2.62}$$

where $\exp\left(-\hat{k}\right)$ is the orbital-rotation operator, which carries unitary transformations of the orbitals. Except for this operator, the parametrization of the model is identical to that of the CI model. In fact, one can resort to the orbital splitting in CASCI or RASCI schemes; by doing so, the corresponding wave functions contain CI coefficients and orbitals variationally optimized under the restrictions imposed by the definition of the active space in each case. The wave functions are constructed by variationally optimizing the expectation value of the energy

$$E = \min_{\mathbf{k},\mathbf{C}} \frac{\langle \mathbf{k},\mathbf{C}|\hat{H}|\mathbf{k},\mathbf{C}\rangle}{\langle \mathbf{k},\mathbf{C}|\mathbf{k},\mathbf{C}\rangle}, \tag{2.63}$$

with respect to the parameters of $\hat{k}$ and the CI coefficients. Here, the problem of optimization is highly non-linear with potentially large number of local solutions and strong couplings among the degrees of freedom. Other drawbacks encountered in MCSCF approaches are related with some amount of arbitrariness, since, as a rule of thumb, an incorrect, unbalanced or ill-defined active space will yield unsatisfactory results [129]. Additionally, the importance of using a good initial set of orbitals for the efficient optimization of MCSCF wave functions, is a common issue when dealing with practical situations [130]. For example, it has been shown that orbitals obtained by diagonalizing the MP2 density matrix constitute a better initial set of orbitals than the canonical Hartree-Fock orbitals, where the virtual orbitals with the lowest orbital energies are too diffuse to describe electron correlation (see section 12.5.5 from Ref. [17]).

Despite these potential problems, orbital optimization in MCSCF approaches ensures the use of adequate orbitals employed for the description of the system, avoiding the introduction of a bias (towards a particular configuration) in the calculations [17]. MCSCF wave functions are flexible in the sense that they can be taken as wave functions *per se* (mostly used for qualitative descriptions) or as orbital generators for more elaborate approaches, as for example in multiconfiguration pair-density functional theory (MPDFT) [131].

Among the most widely used and successful MCSCF methods, is the CASSCF approach [58]. While MCSCF designs a general framework, CASSCF is a specific implementation of MCSCF with a complete active space. Similar to CASCI, in the CASSCF wave function, the inactive orbitals are doubly occupied in all configurations, the active orbitals are subject to no restrictions on their occupations and the secondary orbitals are unoccupied in all configurations. For cases in which electronic excited states are requested, a procedure known as state average (SA) is often employed. In SA-CASSCF, the orbitals are optimized considering a densitry matrix average of all states included, giving a more balanced set

of orbitals for the evaluation of relative energies [122]. Both CASSCF and its SA version, yield a set of orbitals and a reference space, and have shown to be adequate choices for cases where nondynamical correlation plays a significant role, also, CASSCF satisfied the Hellmann-Feynman theorem, while CASCI doesn't. The success of these methods can be attributed to improvements in the efficiency of fast algorithms to solve the full/CAS CI problem [5, 6, 132, 133, 134], second-order orbital optimization [135, 136], and thoughtful approximations to the electron repulsion integral (ERI) tensor [137, 138, 139].

2.4 Dynamic correlation in CI approaches

Nowadays, there exist a plethora of different strategies for a balanced treatment of nondynamic electron correlation. Ranging from relatively simple ideas such as those involved in multireference methods, to more complex situations involving alternative approaches: GW approximation [140, 141], natural orbital functionals [142, 143, 72], geminal-based methods [144], use of transcorrelated Hamiltonians and algorithms such as DMRG [145, 146, 147, 148].

Particularly for variational CAS approaches, despite their relative success in treating NDC, they face limitations for cases in which DC becomes relevant for the accurate description of the system. As early mentioned in section 1.2, the CI approach to the correlation problem imposes a strong computational burden, since a large number of contributions are needed for a correct description of DC. Large CI expansions become slow and computationally intractable, or implicitly, introduce the need for designing elaborate algorithms to handle the required data for calculations. DC treatment is difficult even at MCSSCF level, since it requires a very large number of orbitals with small occupation numbers. In other words, there is no advantage in optimizing the orbitals required for DC.

To overcome these limitations, one may opt to truncate the expansion as in the singles-and-doubles model (CISD, or MRCISD), were the expansion's configurations are generated by carrying out single and double excitations from the full set of reference configurations. However, this leads to further shortcomings such as lack of size consistency and, when breaking spin symmetry, to spin contamination. Also, higher-order excitations may be important for accurately describing certain electronic processes. An unbalanced or lack of correlation treatment might result in the emergence of non-physical solutions and introduces sources of error [66, 83].

Alternatively, approaches which have shown promising improvement, are focused on the introduction of second order perturbative corrections on top of the single reference CAS energies [149, 150, 59, 151]. However, these do not avoid inherit difficulties, such as the "intruder state problem" from perturbation theory. To address these challenges, "n-electron valence state perturbation theory" (NEVPT2) has been applied to CASCI or CASSCF wave functions [152, 153, 154], presenting a novel approach that ensures size consistency, absence of intruder states, spin completeness, and offers an equivalent level of accuracy for both ground and excited states.

Other studies have shown the performance and accuracy achieved through the use of different schemes for computing orbitals to be employed in CASCI. For example, the unrestricted natural orbital (UNO-)CASCI approach [155], which uses the natural orbitals derived from unrestricted HF calculations as the MO basis for a CASCI calculation. While the lack of DC remains, similar approaches to UNO-CASCI have demonstrated to be alternatives delivering correct solutions for specific needs and problems [121, 156, 155] and, in some cases, can provided accurate reference for subsequent treatment of DC [157, 158].

The main objective of this thesis is closely related to the ongoing efforts in the field. It is mainly dedicated to exploring the performance of mingled (WFT/DFT) methodologies [73], where DC is introduced from a density functional approximation (DFA, see below), or provided by an effective potential, into both SR- and MR-CI approaches.

2.5 Density functional theory

In 1964, Pierre Hohenberg and Walter Kohn (HK) [8] demonstrated that for molecules with a non-degenerate ground state, the molecular energy of the ground state, the wave function, and all other electronic properties are uniquely determined by the electronic density of that state (commonly denoted by $\rho_0(\vec{\boldsymbol{r}})$).

The first Hohenberg-Kohn theorem states that the ground state energy of any interacting many-particle system with a given interparticle interaction $\hat{W}$ is uniquely determined by its ground state density. Assuming a non-degenerate ground state, we can express the ground state energy as

$$E_v[\rho] = \langle \Psi[\rho(\vec{\boldsymbol{r}})] | \hat{H} | \Psi[\rho(\vec{\boldsymbol{r}})] \rangle = \int d\vec{\boldsymbol{r}}\, \rho(\vec{\boldsymbol{r}}) v(\vec{\boldsymbol{r}}) + F_{HK}[\rho(\vec{\boldsymbol{r}})], \tag{2.64}$$

where v is used as a label denoting the interacting system, $v(\vec{\boldsymbol{r}})$ is the external potential, ρ the ground state density and F_{HK} the Hohenbeg Kohn functional defined as

$$F_{HK} = \langle \Psi[\rho(\vec{\boldsymbol{r}})] | \hat{T} + \hat{W} | \Psi[\rho(\vec{\boldsymbol{r}})] \rangle \,, \tag{2.65}$$

where $\hat{T}$ and $\hat{W}$ are respectively the kinetic energy and two-electron interaction operators. F_{HK} is a universal functional in the sense that it does not depend on the external potential and can be calculated for all ground state densities of an arbitrary interacting system. Consequently, once the exact form of F_{HK} is known, we can apply it to an arbitrary electronic system, ranging from atoms and molecules to solids [108].

The second theorem provides a variational principle for DFT, analogous to the variational theorem in conventional wave function methods. It states that for a well-behaved trial electronic density, the functional energy yields an energy that is greater than or equal to the exact energy of the system's ground state ($E[\rho_{trial}(\vec{\boldsymbol{r}})] \geq E_0[\rho_0(\vec{\boldsymbol{r}})]$), where the exact energy of the electronic ground state minimizes the energy functional. The energy functional $E_v[\rho]$ satisfies the variational property

$$\begin{aligned} E_0 = E_{v0}[\rho_0] &= \langle \Psi[\rho_0] | \hat{T} + \hat{V}_0 + \hat{W} | \Psi[\rho_0] \rangle \\ &\leq \langle \Psi[\rho_{trial}] | \hat{T} + \hat{V}_0 + \hat{W} | \Psi[\rho_{trial}] \rangle = E_{v0}[\rho_{trial}], \end{aligned} \tag{2.66}$$

where $\hat{V}_0$ is the external potential operator ($\langle \Psi | \hat{V}_0 | \Psi \rangle = \int d\vec{\boldsymbol{r}} \rho(\vec{\boldsymbol{r}}) v_0(\vec{\boldsymbol{r}})$). An alternative form fof the ground state energy expression can be written by explicit splitting of $W[\rho_0]$ into $J[\rho_0]$, the Coulomb energy term (resembling Eq. 1.4), and $E_{xc}[\rho_0]$ the exchange correlation term,

$$E_{v0} = T[\rho_0] + V_0[\rho_0] + J[\rho_0] + E_{xc}[\rho_0], \tag{2.67}$$

with

$$J[\rho_0] = \frac{1}{2}\int d\vec{\boldsymbol{r}} d\vec{\boldsymbol{r}}' \, \frac{\rho_0(\vec{\boldsymbol{r}})\rho_0(\vec{\boldsymbol{r}}')}{|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'|} \tag{2.68}$$

Based on the HK theorems, it is established that the exact energy of the electronic ground state and its density can be calculated without directly resorting to the Schrödinger equation.

2.5.1 Kohn-Sham Density Functional Theory

For the calculation of the electronic energy of the ground state, Kohn and Sham introduced a practical computational scheme for electronic system properties in 1965 [9]. They postulated the existence of a reference system of N non-interacting electrons, with Hamiltonian $\hat{H}_s$ and ground state density $\rho(\vec{\boldsymbol{r}})$, having the same electron density of the interaction system, moving in an external potential v_s. This system is represented by the ground state Slater determinant of orbitals ϕ_p satisfying the following equations (neglecting spin)

$$\left[-\frac{1}{2}\nabla^2 + v_s(\vec{\boldsymbol{r}})\right]\phi_p(\vec{\boldsymbol{r}}) = \epsilon_p \phi_p(\vec{\boldsymbol{r}}), \tag{2.69}$$

with

$$\sum_p^N |\phi_p(\vec{\boldsymbol{r}})|^2 = \rho(\vec{\boldsymbol{r}}). \tag{2.70}$$

By applying the first HK theorem, there must be at most and at least one external potential v_s, to within a constant generating the density ρ. The total energy for this non-interacting system is given by

$$E_{v_s}[\rho] = T_s[\rho] + \int d\vec{\boldsymbol{r}} \, \rho(\vec{\boldsymbol{r}}) v_s(\vec{\boldsymbol{r}}) \tag{2.71}$$

The functional $T_s[\rho]$ is the Kohn-Sham kinetic energy of the non-interacting system, and can be seen as the limit case of a F_{HK} for which inter-particle interactions are neglected. This energy is given by

$$T_s[\rho] = -\frac{1}{2}\sum_p^N \int d\vec{\boldsymbol{r}} \, \phi_p^*(\vec{\boldsymbol{r}})\nabla^2\phi_p(\vec{\boldsymbol{r}}). \tag{2.72}$$

Note that, under the constraint that density must integrate to the total number of electrons, the ground state density can be obtained by the solution of the Euler-Lagrange equations,

$$0 = \frac{\delta}{\delta\rho(\vec{\boldsymbol{r}})}\left[E_{\nu_s}[\rho] - \mu_s \int d\vec{\boldsymbol{r}} \, \rho(\vec{\boldsymbol{r}})\right] = \frac{\delta T_s}{\delta\rho(\vec{\boldsymbol{r}})}[\rho] + v_s(\vec{\boldsymbol{r}}) - \mu_s, \tag{2.73}$$

where μ_s is a Lagrange multiplier introduced to satisfy the previous constraint. Recalling Eq. 2.64 for the interacting system, a comparison with Eqs. 2.71 and 2.72 aims to set the definition of the exchange-correlation energy functional as

$$E_{xc}[\rho] = F_{HK} - \frac{1}{2}\int d\vec{\boldsymbol{r}}_1 d\vec{\boldsymbol{r}}_2 \, \frac{\rho(\vec{\boldsymbol{r}}_1)\rho(\vec{\boldsymbol{r}}_2)}{|\vec{\boldsymbol{r}}_1 - \vec{\boldsymbol{r}}_2|} - T_s[\rho] \tag{2.74}$$

As the F_{HK} functional is only defined for ground state interacting systems, and T_s for ground state densities of non-interacting systems, until here, we have operated on the premise that every interacting system possesses an equivalent non-interacting system that yields an identical ground state density, this was conclusively established by Kohn and Sham [9]. With this foundational assumption recognized, the Euler-Lagrange equations for the interacting system can be expressed as follows

$$\begin{aligned} 0 &= \frac{\delta}{\delta\rho(\vec{\boldsymbol{r}})}\left[E_\nu[\rho] - \mu_s \int d\vec{\boldsymbol{r}}\, \rho(\vec{\boldsymbol{r}})\right] \\ &= \frac{\delta T_s}{\delta\rho(\vec{\boldsymbol{r}})}[\rho] + v(\vec{\boldsymbol{r}}) + \int d\vec{\boldsymbol{r}}' \, \frac{\rho(\vec{\boldsymbol{r}}')}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} + v_{xc}([\rho];\vec{\boldsymbol{r}}) - \mu_s, \end{aligned} \tag{2.75}$$

where the exchange-correlation potential v_{xc} is defined as

$$v_{xc}([\rho];\vec{\boldsymbol{r}}) = \frac{\delta E_{xc}}{\delta\rho(\vec{\boldsymbol{r}})}[\rho(\vec{\boldsymbol{r}})] \tag{2.76}$$

Like HF, Kohn-Sham DFT (KS-DFT) is an effective non-interacting particle model with a Coulomb potential plus a so-called exchange-correlation functional, which combines exchange and correlation effects into a single term. In contrast to variational WFT methodologies, KS-DFT offers a straightforward and computationally efficient method to incorporate DC by explicitly accounting for electron correlation in the evaluation of this functional. Using the fact that both, the noninteracting and interacting systems are solved for the same density, then

$$v_s(\vec{\boldsymbol{r}}) = v(\vec{\boldsymbol{r}}) + \int d\vec{\boldsymbol{r}}' \, \frac{\rho(\vec{\boldsymbol{r}}')}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} + v_{xc}([\rho];\vec{\boldsymbol{r}}) \tag{2.77}$$

to within a constant (that can be chosen such that $v_{xc} \to 0$ when $|r| \to \infty$). The density for the interacting system can be obtained by solving the KS equations (with density given as in Eq. 2.70),

$$\left[-\frac{1}{2}\nabla^2 + v(\vec{\boldsymbol{r}}) + \int d\vec{\boldsymbol{r}}' \, \frac{\rho(\vec{\boldsymbol{r}}')}{\left|\vec{\boldsymbol{r}} - \vec{\boldsymbol{r}}'\right|} + v_{xc}([\rho];\vec{\boldsymbol{r}})\right]\phi_p(\vec{\boldsymbol{r}}) = \epsilon_p\phi_p(\vec{\boldsymbol{r}}). \tag{2.78}$$

With this, the problem is simplified to that of finding good approximations to the F_{HK} implicit in $v_{xc}([\rho];\vec{\boldsymbol{r}})$, the key of KS-DFT is the introduction of correlation energy (mostly dynamic) through the solutions of a SCF-like set of equations, the cos of which has a scaling similar to HF ($\mathcal{O}(N^4)$) or lower ($\mathcal{O}(N^3)$ or $\mathcal{O}(N^2)$). This allows DFT to handle a wider range of systems, including those with significant electron correlation, like transition metals and molecules with stronger electron interactions [159]. This is in principle true, but in practice such calculations are limited by the approximations made to the exchange-correlation functional, the multiconfigurational nature of the electronic states, and to spurious energy terms (such those appearing due to self-interaction) [160].

2.6 Overview of selected methods for simultaneous treating of DC and NDC

Although there is a plethora of methods able to efficiently recover either NDC or DC, very few approaches are able to treat both kinds of correlation in a balanced manner. This section briefly summarizes the main features of some current approaches designed for this purpose. Essentials on illustrative approaches with different ways to treat DC are considered here. We consider to include those using a perturbative correction, through DFT in mingled WFT+DFT methods and others involving the adiabatic connection formula.

2.6.1 CASPT2 and CAS-NEVPT2

Here, the basic CASPT2 theory is illustrated with discussions on the most serious problem of the approach, the intruder state problem. Original CASPT2 theory was cast in light of four constraints [161]:

- The use of Rayleigh-Schrödinger perturbation theory (RSPT) and the intermediate normalization approach.
- The unperturbed reference state should be the CASSCF wave function.
- The reference Hamiltonian, $\hat{H}_0^{CASPT2}$, should be an effective one-electron operator.
- The CASPT2 should reduce to the MP2 formalism in the case the CASSCF CI expansion is reduced to a single Slater determinant.

The central idea here is to correct the CASSCF wave function and energy via perturbation theory. To this end, it is necessary to define a zeroth order Hamiltonian that must be relatively simple, and compatible with the previous constraints. This Hamiltonian can be written in terms of a generalized Fock operator (Eq. 2.48). In order to show a suitable form of $\hat{H}_0^{CASPT2}$, consider the restricted closed-shell Fock operator (Eq. 2.8) written in chemist's notation as

$$\hat{F} = \sum_{pq} \hat{E}_{pq} \left(h_{pq} + \sum_{k} \left[2(pq|kk) - (pk|qk) \right] \right) \tag{2.79}$$

where the summation is over the occupied SCF orbitals, $\hat{E}_{pq} = \sum_{\sigma} \hat{\psi}^{\dagger}_{p\sigma} \hat{\psi}_{q\sigma}$ is the spin-traced excitation operator, the matrix elements h_{pq} and $(pq|rs)$ are the one-electron integrals (containing the kinetic energy and the electron-nuclear attraction) and the electron-electron repulsion integrals, respectively. The generalized Fock operator for a CASPT2 theory, written in this notation is

$$\hat{F} = \sum_{pq} f_{pq} \hat{E}_{pq} \tag{2.80}$$

with

$$f_{pq} = h_{pq} + \sum_{rs} {}^{1}D_{rs} \left[(pq|rs) - \frac{1}{2}(pr|qs) \right], \tag{2.81}$$

where ${}^1D_{rs}$ are elements of the 1-RDM. $\hat{F}$ does not have CASSCF wave functions as eigenfunctions. We can, however, define an alternative Hamiltonian with CASSCF wave functions as eigenfunctions, namely [162]

$$\hat{H}_0^{CASPT2} = \sum_I |\Psi_I\rangle \langle\Psi_I|\hat{F}|\Psi_I\rangle \langle\Psi_I| + \sum_{K,L} |\Psi_K\rangle \langle\Psi_K|\hat{F}|\Psi_L\rangle \langle\Psi_L| . \tag{2.82}$$

Where Ψ_I is the I-th root of the CASSCF procedure and Ψ_K (and Ψ_L) are external electronic configurations not found in the CI space of the CASSCF expansion. The next step is to compute the first-order correction to the wave function ($\left|\Psi^{(1)}\right\rangle$), which consists of all the configurations that can interact directly with the zeroth-order wave function through the Hamiltonian. This space is referred to as the first-order interacting space (FOIS). To identify and manipulate FOIS, the configuration space is commonly separated into four subspaces chosen by convenience to be, V_0, the one-dimensional space of the reference CASSCF wave function, V_P, containing all other CASSCF CI expansions consistent with the number of active electrons, active orbitals and spin multiplicity as the reference wave function. The subspace V_{SD}, is spanned by all possible single and double electron replacements generated from the reference CASSCF wave function, and V_{TQ} corresponds to the complement of the first thee subspaces. Considering whit, the reference Hamiltonian from Eq. 2.82 can be expressed as

$$\begin{aligned}\hat{H}_0^{CASPT2} = &\left|\Psi^{(0)}\right\rangle \left\langle\Psi^{(0)}\right|\hat{F}\left|\Psi^{(0)}\right\rangle \left\langle\Psi^{(0)}\right| \\ &+ \sum_{K,L\in P} |\Psi_K\rangle \langle\Psi_K|\hat{F}|\Psi_L\rangle \langle\Psi_L| \\ &+ \sum_{K,L\in SD} |\Psi_K\rangle \langle\Psi_K|\hat{F}|\Psi_L\rangle \langle\Psi_L| \\ &+ \sum_{K,L\in TQ} |\Psi_K\rangle \langle\Psi_K|\hat{F}|\Psi_L\rangle \langle\Psi_L| ,\end{aligned} \tag{2.83}$$

only those configurations in V_{SD} will interact with the reference wave function via the total Hamiltonian, and will constitute the FOIS. The interaction coming from those configurations in V_P is neglected due to the variational principle (V_P is null in MP2). Therefore $\left|\Psi^{(1)}\right\rangle$ can be expressed as a linear combination of the configurations, $|\Psi_J\rangle$, belonging to the SD subspace,

$$\left|\Psi^{(1)}\right\rangle = \sum_{J\in SD} C_J |\Psi_J\rangle . \tag{2.84}$$

Being $\hat{V}$ a perturbation term accounting for the remaining part of the electron-electron interaction not accounted by $\hat{H}_0^{CASPT2}$,

$$\mathbf{H}(z) = \mathbf{H}_0 + z\mathbf{V}. \tag{2.85}$$

In doing so, the coefficients can be determined by solving the system of equations,

$$\sum_{J\in SD} C_J \langle\Psi_I|\hat{H}_0 - E_0|\Psi_J\rangle = -\left\langle\Psi_I\right|\hat{V}\left|\Psi^{(0)}\right\rangle , \quad I = 1,\ldots,M, \quad I \in SD \tag{2.86}$$

which can be put into a matrix-vector notation,

$$(\mathbf{H}_0 - E_0 I)\mathbf{C} = -\mathbf{V}, \tag{2.87}$$

where I is the identity matrix. The second-order correction to the energy is given by

$$E^{(2)} = \left\langle \Psi^{(0)} \middle| \hat{V} \middle| \Psi^{(1)} \right\rangle = \mathbf{V}^{\dagger}\mathbf{C} \tag{2.88}$$

Here, a problem arises, since not all the configurations within V_{SD}, as previously defined, are necessarily linearly independent. Therefore an intermediary step, involving diagonalization of the overlap matrix $\mathbf{S}_{IJ} = \langle \Psi_I | \Psi_J \rangle$ and elimination of those eigenvectors with zero or very small eigenvalues, is needed in order to remove the linear and near-linear dependences among configurations [163].

A more serious problem is that CASPT2 suffers from an issue common to other MRPT-based approaches, the intruder-state problem. Provided the orthogonality of the eigenfunctions of $\hat{H}_0$, an analytical expression for the coefficients C_I is simply

$$C_I = -\frac{\langle \Psi_I | \hat{V} | \Psi_0 \rangle}{\epsilon_I - E_0} \tag{2.89}$$

where ϵ_I is the corresponding energy eigenvalue from the equation $\hat{H}_0 \Psi_I = \epsilon_I \Psi_I$. The intruder-state problem arises when the energy denominator in Eq. 2.89 vanishes, leading to an infinitely large coefficient. This is promoted by a degeneration (or near-degeneration) between the states $|\Psi_0\rangle$ and the intruder Ψ_I.

A simple solution to address this issue involves incorporating the intruder state $|\Psi_I\rangle$ into the reference wave function. However, in CASPT2, this is often impractical or undesirable due to its impact on the active space during CASSCF optimization. This alteration may lead to the emergence of new intruder states, necessitating additional electronic states in state-averaging procedures, or resulting in prohibitively expensive calculations [164].

An alternative solution is to modify the partitioning of the Hamiltonian. If the degeneracy between the intruder state $|\Psi_I\rangle$ and $|\Psi_0\rangle$ is a consequence of the approximate description by the zeroth-order Hamiltonian, one can adjust the structure of $\hat{H}_0^{CASPT2}$ to eliminate this accidental degeneracy. Strategies such as the use of a Dyall's Hamiltonian [165] or the introduction of real and imaginary level shift techniques has proven effective in this regard [166, 167].

The CAS N-electron valence state second-order perturbation theory (NEVPT2) serves as an alternative to CASPT2. NEVPT2 is a second-order perturbation theory with a CAS (or a CAS-like) reference wave function originally developed by Angeli and co-workers [168, 152, 153, 154]. In contrast to CASPT2, it utilizes a Dyall Hamiltonian [165] as the zeroth-order Hamiltonian, given by

$$\hat{H}_0^{Dyall} = \sum_i \epsilon_i \hat{E}_{ii} + \sum_{tu} h'_{tu} \hat{E}_{tu} + \frac{1}{2} \sum_{tuvw} (tu|vw) \hat{e}_{tu,vw} + \sum_a \epsilon_a \hat{E}_{aa}, \tag{2.90}$$

where

$$\sum_{pq} h'_{pq} = h_{pq} + \sum_{i} [2(pq|ii) - (pi|qi)] \tag{2.91}$$

and

$$\hat{e}_{pq,rs} = \hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps} \tag{2.92}$$

Operator (2.90) is invariant to rotations of orbitals within the active space, commutes with the $\hat{S}^2$ operator, and yields the CASSCF energy (E_{CASSCF}) when applied to the CASSCF reference wave function. By design, the choice of this full active space Hamiltonian effectively addresses the problem of intruder states (unless the active space is poorly defined), making CAS-NEVPT2 virtually intruder-state-free [169].

Various contraction schemes to span the first-order interacting space exist, including strongly contracted (SC), internally contracted (IC), and uncontracted (UC) constructions [168]. In any of these schemes, the FOIS can be grouped into eight distinct excitation subspaces, depending on the number of hole and particle labels in the excitation operators. Among them, SC-NEVPT2 is the most efficient, as it completely avoids any iterative procedure for the first-order wave function or the storage of amplitudes. However, it lacks unitary invariance concerning rotations within the doubly occupied and virtual spaces.

Recently, a benchmark study [170] evaluated the utility and accuracy of CASPT2, NEVPT2, and other single-reference WFT methods, involving the computation of vertical transition energies for over 250 electronic states in 35 small to medium-sized organic molecules. The global statistical analyses reveal that standard CASPT2 performs poorly, especially for Rydberg transitions. NEVPT2 shows small overestimations of vertical transition energies, providing more accurate estimates for different subsets of electronic states. Among other comparisons, results from the state-average version of CASSCF, neglecting dynamic correlation, are detrimental, while NEVPT2 only slightly deteriorates results. Delivering similar accuracies as the second-order coupled-cluster (CC2) model and CCSD.

CASPT2 and NEVPT2 are valuable methodologies, known for their strength in simultaneously handling NDC and DC for arbitrary electronic states, at a reasonable computational cost, providing exact spin states.However, a notable limitation is that they are not black box methods and require considerable experience to be used effectively. Additionally, for accurate results, the active space must generally be larger than the minimum required to describe only the NDC.

2.6.2 CAS-DFT

In revisiting WFT-DFT mingled approaches to develop a CASSCF combination with DFT, Savin and co-workers devised two strategies. The first relies on the use of the on-top pair density [171], the second, based on a correlation functional whose magnitude should consider the correlation already present in CASSCF [172]. The new correlation functional was developed from the free-electron gas model, where the exchange-correlation energy per particle ϵ_c becomes a functional of the density ($\rho(\vec{r})$), its gradient (in the case of GGA versions) and the on-top pair density ($\Pi(\vec{r})$), and is scaled through a control parameter function (φ) encoding the reduction of the correlation energy when part of the correlation energy is calculated separately by CASSCF,

$$\epsilon_c\left(\rho(\vec{\boldsymbol{r}}), \boldsymbol{\nabla}\rho(\vec{\boldsymbol{r}}), \Pi(\vec{\boldsymbol{r}}), \frac{k_g}{k_F}\right) = \varphi\left(r_s, \frac{k_g}{k_F}\right)\epsilon_c^{\text{std}}(\rho, \boldsymbol{\nabla}\rho, \Pi) \tag{2.93}$$

where r_s is the Wigner-Seitz radius ($r_s = (3/4\pi\rho)^{1/3}$), $\epsilon_c^{\text{std}}(\rho, \boldsymbol{\nabla}\rho, \Pi)$ is the correlation energy given as in the standard LDA or GGA formalism. The factor $\varphi(r_s, \frac{k_g}{k_F})$, was established from parametrization to free-electron gas model, such that the following conditions are satisfied: $\varphi \to 1$ when $k_g \to k_F$ and $\varphi \to 0$ when $k_g \to \infty$ [172]. The total energy in the original CAS-DFT model is

$$E_{CAS-DFT} = E_{CAS} + E^{DFT}[\rho_{CAS}, \Pi_{CAS}, \varphi], \tag{2.94}$$

for which the DFT contribution, uses ϵ_c from Eq. 2.93. It is crucial for the control parameter, φ, to meet specific theoretical criteria:

1. It should possess physical significance not only for the system under consideration, but also for the electron gas.
2. It must establish a suitable distinction between static and DC for every density, eliminating artifacts (such as steps in the potential curve) in the reference CI calculation.

With this strategy, the CASSCF energy can be improved by adding a weighted DC delivered by the DFA in use. Originally, Savin and co-workers used this approach for small systems: to compute correlation energies for the isoelectronic series of He and Be and the potential energy curve of H_2. They concluded that introducing a correction factor (depending on the active space), enables a smooth transition between the limiting case of a pure density-functional and a full-CI calculation.

This scheme was further modified by considering the DC delivered by the active space [75]. For this purpose, to avoid double counting of DC, the DFT energy is scaled by an additional factor f (were $0 \leq f \leq 1$). To establish an indicator for the size of the active space, a reference density was defined as

$$\rho_{ref}(\vec{\boldsymbol{r}}) = \sum_{p}^{occ} 2|\phi_p(\vec{\boldsymbol{r}})|^2, \tag{2.95}$$

where the sum runs over all occupied orbitals, and both inactive and active orbitals are considered to be doubly occupied. The ratio $\rho_{ref}(\vec{\boldsymbol{r}})/\rho(\vec{\boldsymbol{r}})$ (where $\rho(\vec{\boldsymbol{r}})$ is the total density) is greater than or equal to 1, and increases with the size of the active space, i.e. the smaller the amount of the DC energy to be covered by $E_C^{CAS-DFT}$ becomes.

The calculation of f involves the following steps. Initially, the correlation energy per particle, denoted as $\epsilon_c(\rho_{ref}(\vec{\boldsymbol{r}}), \rho(\vec{\boldsymbol{r}}))$ is calculated using the reference density as the density of the active space. Following this, all virtual orbitals are included in the active space, and the correlation energy is recalculated as $\epsilon_c(\rho_{ref}(\vec{\boldsymbol{r}}), \infty)$. The scaling factor is then determined as

$$f = 1 - \frac{\epsilon_c(\rho_{ref}(\vec{\boldsymbol{r}}), \rho(\vec{\boldsymbol{r}}))}{\epsilon_c(\rho_{ref}(\vec{\boldsymbol{r}}), \infty)}, \tag{2.96}$$

CAS-DFT provides a description for both static and DC effects at moderate computational costs, representing an initial starting point for mingled schemes of this type. But the method faced various problems. So far in the analysis, the theory assumes that all occupied orbitals are in the active space, but in most conventional CASSCF calculations, there are also inactive orbitals, and with such orbitals there can be also inconsistencies in the model. Later, it was realized that the scaling factor introduced in this way did not entirely remove double counting effects [74].

Based on these approaches, additional CASSCF or CASCI with DFT [173, 174] were introduced, mostly, to provide a proper description of bond dissociation processes in the electronic ground state, without focusing on electronically excited states. Related approaches incorporate elements such as natural orbitals [175], perturbation theory corrections [176] or self-consistent optimization of CI and MO coefficients [177], allowing to expand the method to excited states studies.

2.6.3 DFT/MRCI

The DFT/MRCI method appears to be a highly accurate option able to compute excitation energies compared to experimental values [178], while it can be considered a WFT+DFT method, its foundations are different from other approaches such as those based on the range-separated formalism. DFT/MRCI was originally devised for the description of spectral properties of singlet and triplet excited states [179], and has been extended to a more general multireference method applicable to states of all spin multiplicities. Central in this development was the aim to provide a general-purpose method for computing electronic spectra and excited state properties. DFT/MRCI is intended to be of "black-box" character, that is, applied straihtforwardly starting from a ground state SCF solution without previous knowledge of the system's attributes. In contrast with methods from the CAS family or CASPT2 variants, in DFT/MRCI the reference space is automatically adjusted to the requierements of the given problem [178].

In DFT/MRCI, KS orbitals are utilized as the one-electron basis from which the configuration space is constructed. This choice was motivated by the desire to generate a more compact MRCI expansion than in the basis of HF orbitals. It also incorporates KS orbital energies into the Hamiltonian. DC is well accounted for by DFT whereas HF-based MRCI needs long expansions of doubly or even higher excited configurations to capture this effect. In contrast, MRCI is well suited to describe static electron correlation, which arises from near-degeneracies and is not covered by KS-DFT.

While the main emphasis of this thesis is on single-reference approaches, the subsequent discussion is devoted to showing the essential features of the DFT/MRCI method. With this aim, a slight change of notation will be introduced for this section. For DFT/MRCI it is adequate to represent the Hamiltonian in a basis of CSFs, $|\mathrm{v}\omega\rangle$, characterized by v, a spatial occupation number vector and ω, a spin-coupling pattern pertaining to the open shells in v. In absence of spin dependent operators, the Hamiltonian matrix can be written in a block-diagonal form for each multiplicity. MRCI can thus be performed separately for singlets, doublets, triplets, etc.

The total set of CSFs $\Omega = \{|\mathrm{v}\,\omega\rangle\}$ is obtained from the set of single and double excitations from a small reference space $\Omega_0 = \{|\mathrm{v}_0\,\omega\rangle\}$ (here v_0 denotes the reference space configurations). The on-diagonal elements of the DFT/MRCI Hamiltonian matrix take the following form

$$\langle \mathrm{v}\omega|\hat{H}^{DFT} - E_{DFT}|\mathrm{v}\omega\rangle = \langle \mathrm{v}\omega|\hat{H} - E_{SCF}|\mathrm{v}\omega\rangle + \sum_p \Delta \mathrm{v}_p(\epsilon_p^{KS} - F_{pp}) + \Delta E_x + \Delta E_c, \tag{2.97}$$

which resembles the form of the sum of the exact result plus DFT corrections. Here, $\Delta \mathrm{v}_p = \mathrm{v}_p - \bar{\mathrm{v}}_p$ is the difference of the occupation of the pth spatial orbital relative to a base occupation $\bar{\mathrm{v}}_p$, chosen as the HF occupation. The quantities ϵ_p^{KS} and F_{pp} denote, respectively, the KS orbital energies and diagonal elements of the Fock operator in the KS orbital basis. Finally, ΔE_x and ΔE_c are Coulomb and exchange corrections.

The role played by DFT corrections is to account for DC that in an *ab initio* CI treatment, would be described by (double and higher) excitations into energetically high-lying orbitals. Thus, to avoid double counting in DFT/MRCI, the off-diagonal Hamiltonian matrix elements are modified. This is achieved by excluding high energy orbitals from the treatment. A damping function for off-diagonal matrix elements is introduced in the original DFT/MRCI, which depends on the energetic separation of the bra and ket CSFs,

$$\left\langle \mathrm{v}\omega\middle|\hat{H}^{DFT} - E_{DFT}\middle|\mathrm{v}'\omega'\right\rangle = \left\langle \mathrm{v}\omega\middle|\hat{H} - E_{SCF}\middle|\mathrm{v}'\omega'\right\rangle \cdot D(\Delta E_{\mathrm{vv}'}), \tag{2.98}$$

where the energy difference is

$$\Delta E_{\mathrm{vv}'} = \frac{1}{n_\omega}\sum_{\omega}^{n_\omega} H^{DFT}_{\mathrm{v}\omega,\mathrm{v}\omega} - \frac{1}{n'_\omega}\sum_{\omega'}^{n'_\omega} H^{DFT}_{\mathrm{v}'\omega',\mathrm{v}'\omega'}. \tag{2.99}$$

The latter term denotes the spin-coupling-averaged difference between the diagonal elements corresponding to the spatial occupations w and w', being n_ω and $n_{\omega'}$, the total number of spin-coupling patterns in v, v′, respectively. In practice, the function $D(\Delta E_{vv'})$ is typically chosen to be an exponential [179, 180] or an inverse arctangent function [181, 182], for which some empirical parameters are introduced. Damping the off-diagonal terms has the effect of minimizing the double counting of DC. This also implies that CSFs from the reference space become decoupled from the vast majority of CSFs spanning the FOIS. This ability to discard almost all of the FOIS CSFs is one of the method's most important features. In practice, this is achieved using an energy-based selection criterion in addition to the previously mentioned damping. For each FOIS configuration w, the quantity

$$d_w = \sum_p \Delta w_p \epsilon_p^{KS} - \Delta E_{sel} \tag{2.100}$$

is used, where ΔE_{sel} is conventionally chosen as 0.8 or 1.0 a.u. [179]. If d_w is lower than the highest reference space eigenvalue of interest, then all of the CSFs generated from configuration w are selected for inclusion.

2.6.4 MCPDFT and CASΠDFT methods

While DFT/MRCI takes advantage of the KS orbitals into the WFT framework, another possibility is to take advantage of a multiconfigurational wave function and define a density functionals accounting for the non-classical energy. This strategy is adopted in Multiconfigurational Pair Density Functional Theory (MCPDFT) [171, 183, 131, 184], in which one uses

orbitals from a MCSCF wave function to compute the kinetic energy and classical Coulomb energy of the multiconfigurational wave function. The remaining non-classical terms are computed from a DFA as in DFT.

As indicated in previous sections, a multideterminant situation posses differences from the single determinant case. As long as the wave function is represented by a single determinant, exchange-correlation density functionals can be formulated in terms of individual spin densities.

For a given external potential $v(\vec{\boldsymbol{r}})$ and electron number N, a self-consistent solution for the KS one-electron equations yields a pair of electron densities $\rho_\alpha(\vec{\boldsymbol{r}})$ and $\rho_\beta(\vec{\boldsymbol{r}})$ and a total energy E. In the KS-DFT framework, these quantities are best interpreted as approximations to the the corresponding physical quantities, the total density $\rho(\vec{\boldsymbol{r}}) = \rho_\alpha(\vec{\boldsymbol{r}}) + \rho_\beta(\vec{\boldsymbol{r}})$, and spin magnetization density $m(\vec{\boldsymbol{r}}) = \rho_\alpha(\vec{\boldsymbol{r}}) - \rho_\beta(\vec{\boldsymbol{r}})$.

This standard interpretation faces a symmetry dilemma illustrated by the binding energy curve of H_2 [171]. At equilibrium, KS equations yield a self-consistent ground state, but at larger separations, a second, lower-energy, "broken symmetry" solution emerges. This solution describes two hydrogen atoms with opposite electron spins. Calculated dissociation energies for the solutions pose a dilemma: the LSD or GGA energy is nearly exact for the broken-symmetry solution with incorrect spin densities but deviates significantly when imposing the correct singlet spin symmetry. This symmetry breaking occurs in Hartree-Fock theory as well [185, 16].

In other words, an approximate wave function may sometimes obtain a variational lower energy state by breaking the symmetry that the exact wave function should have. For MC wave functions, spin-DFT features a symmetry dilema and $m(\vec{\boldsymbol{r}})$ is no longer a proper description. To preserve multiplet spin degeneracies, a reformulation of the standard DFAs was advised [171]. This has been one of the main goals of the theoretical development of MCPDFT. In MCPDFT the so-called on-top functionals are used instead, these are formulated in terms of the density, $\rho(\vec{\boldsymbol{r}})$, and the on-top pair density, $\Pi_I(\vec{\boldsymbol{r}})$, the MCPDFT energy is then defined as the sum of the terms computed from the MC wave function plus the on-top functional:

$$E_{MCPDFT} = V_{nuc} + \sum_{pq}\left[h_{pq}\,{}^1D_{pq} + \frac{1}{2}\sum_{rs}(pq|rs)\,{}^1D_{pq}\,{}^1D_{rs}\right] + E_{OT}[\rho, \Pi_I]. \quad (2.101)$$

This expression contains the sum of nuclear-nuclear repulsion (V_{nuc}), the nuclear-electron attraction (contained in h_{pq}), and the electron-electron repulsion (given by $(pq|rs)$ terms). In MCPDFT, the double counting problem is avoided since the energy of the MC wave function is not used, rather the energy is calculated as the sum of a kinetic energy term, a classical Coulomb term and an on-top energy.

Recent developments in MCPDFT [183, 131, 186] utilize a normalized form of the on-top density

$$R(\vec{\boldsymbol{r}}) = \frac{4\Pi(\vec{\boldsymbol{r}})}{|\rho(\vec{\boldsymbol{r}})|^2}, \quad (2.102)$$

which becomes 1 in the closed-shell HF case. Moreover, MCSCF wave functions are spin eigenfunctions with physical spin densities, while current available DFAs use unphysical

effective spin densities to mimic the effects of unpaired electrons. As a consequence, when applied to strongly correlated systems, DFAs are inaccurate when physical spin densities are used. Therefore, a key strategy proposed by Gagliardi and co-workers [183, 131] is to extract effective spin densities from MCSCF wave functions that will allow to use KS-DFAs. The on-top functionals in this case are translated from common exchange-correlation functional in LDA or GGA frameworks by setting

$$\begin{aligned} m(\vec{\boldsymbol{r}}) =& \rho(\vec{\boldsymbol{r}})[1 - R(\vec{\boldsymbol{r}})]^{1/2} \quad \text{if} \quad R \leq 1 \\ m(\vec{\boldsymbol{r}}) =& 0 \quad \text{if} \quad R(\vec{\boldsymbol{r}}) > 1. \end{aligned} \tag{2.103}$$

This would define the on top functional $E_{OT}[\rho(\vec{\boldsymbol{r}}), \Pi(\vec{\boldsymbol{r}})]$ by "translation" or fully translation of a KS exchange-correlation functional that depends on $\rho(\vec{\boldsymbol{r}})$, $m(\vec{\boldsymbol{r}})$ and the magnitude of their gradients, $\rho'(\vec{\boldsymbol{r}})$, $m'(\vec{\boldsymbol{r}})$. The equation for simple translation being

$$E_{OT}[\rho(\vec{\boldsymbol{r}}), \Pi(\vec{\boldsymbol{r}})] = E_{xc}[\rho(\vec{\boldsymbol{r}}), m(\vec{\boldsymbol{r}}), \rho'(\vec{\boldsymbol{r}}), m'(\vec{\boldsymbol{r}})]. \tag{2.104}$$

In principle, a MCPDFT approach may be combined with any wave function method. The most commonly used is CASSCF. In a MCPDFT calculation, the CASSCF reference wave function is computed and used to get the kinetic energy, the density and on-top pair density. Then these functions are used as inputs to obtain the total energy. The quality of the approach relies on two main points: (1) to compute a good quality wave function and (2) to use a good exchange-correlation density functional. Within this in mind, various works have been devoted to combine alternative WFT approaches other than CASSCF, and to develop well behaved on-top functionals [187].

Alternatively, a combined WFT-DFT approach using the on-top pair density has been proposed, the so-called CASΠDFT method [188, 189], in which the NDC is obtained with a CAS wave fucntion, while DC is calculated with a (corrected) DFT energy functional. The new feature introduced here is that this modified correlation functional is sensitive to the NDC already included in the WFT part. This is achieved through a correction function that involves the on-top pair density and locally adjusts the DC energy to account for suppression of DC (SDC) and enhancement of DC (EDC).

A local quantity that conveys information about electron correlation from WFT to DFT is the conditional density ρ_{cond}, defined as the ratio of the on-top pair density to the electron density, and defined as

$$\rho_{\text{cond}}(\vec{\boldsymbol{r}}) = \frac{\Pi(\vec{\boldsymbol{r}})}{\rho(\vec{\boldsymbol{r}})}, \tag{2.105}$$

which yields the electron density at $\vec{\boldsymbol{r}}$ on condition that a reference electron occupies the same position.

The conditional density for a closed-shell single determinant wave function is $\rho_{\text{cond}}(\vec{\boldsymbol{r}}) = \rho(\vec{\boldsymbol{r}})/2$. In the case of a MC wave function, $\rho_{\text{cond}}[\Psi^{MC}](\vec{\boldsymbol{r}})$ can be less than, equal to, or greater than $\rho(\vec{\boldsymbol{r}})/2$, indicating electron correlation effects. For regions of space where $\rho_{\text{cond}}\Psi^{MC} < \rho(\vec{\boldsymbol{r}})/2$, a depletion of density is manifested. This effect corresponds to the previously mentioned SDC [188]. The opposite case arises when the conditional density increases, $\rho_{\text{cond}}[\Psi^{MC}](\vec{\boldsymbol{r}}) > \rho(\vec{\boldsymbol{r}})/2$. In such cases, the DC energy is also increased, effect referred as EDC.

It has been argued [189] that the ratio,

$$X(\vec{\boldsymbol{r}}) = \frac{\rho_{\text{cond}}(\vec{\boldsymbol{r}})}{\rho(\vec{\boldsymbol{r}})/2} = \frac{2\Pi(\vec{\boldsymbol{r}})}{\rho(\vec{\boldsymbol{r}})^2} \tag{2.106}$$

indicates either, the SDC ($X(\vec{\boldsymbol{r}}) < 1$) or the EDC ($X(\vec{\boldsymbol{r}}) > 1$) effect. In CASΠDFT the total energy is computed as a sum of the CASSCF energy and the DFT-based DC correction as

$$E^{CAS\Pi DFT} = E^{CASSCF} + E_{DC}^{\Pi DFT}[\rho^{CASSCF}, \Pi^{CASSCF}], \tag{2.107}$$

with

$$E_{DC}^{\Pi DFT}[\rho^{CASSCF}, \Pi^{CASSCF}] = \int d\vec{\boldsymbol{r}} P[X](\vec{\boldsymbol{r}})\epsilon_c[\rho^{CASSCF}](\vec{\boldsymbol{r}})], \tag{2.108}$$

where the multiplicative function, $P[X]$, is introduced to adjust the DFT correlation energy to account the SDC and EDC effects. The DFT contribution is evaluated in a post-CAS fashion, i.e. after a CASSCF calculation is performed, the DFT contibution is evaluated and added. The form of $P[X]$ has been proposed as

$$P[X(\vec{\boldsymbol{r}})] = P^{SDC}[X(\vec{\boldsymbol{r}})] = \frac{aX(\vec{\boldsymbol{r}})}{1 + bX(\vec{\boldsymbol{r}})} \quad X \leq 1 \tag{2.109}$$

$$P[X(\vec{\boldsymbol{r}})] = P^{EDC}[X(\vec{\boldsymbol{r}})] = c(X(\vec{\boldsymbol{r}})^{1/4}) - \frac{d(X(\vec{\boldsymbol{r}}) - g)^2}{(1-g)^2} \quad X > 1. \tag{2.110}$$

Ensuring the independent representation of the SDC and EDC regions. The parameters $a - g$ were selected as follows. The P^{SDC} is chosen to provide the correct SDC asymptotics in the $X \to 0$ and $X \to 1$ limits. In the limit $X \to 0$, P^{SDC} vanishes, reflecting strong SDC due to strong NDC. The opposite limit $X \to 1$ represents a weak NDC regime, which does not suppress DC. This condition fixes $b = a - 1$. The single free parameter a is fitted to reproduce the reference data for the ground state of H_2 [189].

The condition $P^{EDC}(1) = 1$ raises from continuity in the definitions 2.109 and 2.110, and fixes one of the three remaining parameters from 2.110,

$$d = \frac{c-1}{(1-g)^2}. \tag{2.111}$$

Finally, parameters c and g are fitted to reproduce the reference data for the excited singlet ${}^1\sigma_u^+$ state of H_2. Within such parametrization it has been shown that correlation energy computed within CASΠDFT closely reproduces the reference H_2 FCI data with chemical accuracy. Suggested values for the free parameters involved in the $P[X]$ functions are, $a = 0.2$, $c = 2.6$ and $g = 1.5$. These values have been successfully used to describe potential energy curves of molecules [190, 191].

The strategy employed by the CASΠDFT method allows for an increased accuracy in vertical gap calculations, attributed to its capability to incorporate DC in a balanced manner for both ground and excited states. Furthermore, the computational effort required for calculating the correlation energy is negligible due to the evaluation of solely local functions. This is in contrast to other WFT-DFT approaches, which, for instance, may necessitate the evaluation of two-electron integrals.

2.6.5 Range separation formalism

As previously mentioned, a way to avoid double counting of the correlation energy is to use the range separation. The formalism proposed by Savin and co-workers [192] is based on the split of the Coulomb interaction operator, r_{12}^{-1}, into short-range (SR) and long-range (LR) components, as

$$\frac{1}{r_{12}} = \hat{v}_{ee}^{SR}(r_{12}) + \hat{v}_{ee}^{LR}(r_{12}). \tag{2.112}$$

This splitting is restricted to the following conditions: for the long-range part, it must be finite at the electron-electron coalescence, $r_{12} \to 0$

$$\lim_{r_{12}\to 0} \hat{v}_{ee}^{LR}(r_{12}) = \quad \text{constant} \quad , \tag{2.113}$$

while at large separation limit, the operator must reduce to the Coulomb interaction

$$\lim_{r_{12}\to \infty} \hat{v}_{ee}^{LR}(r_{12}) = \frac{1}{r_{12}}. \tag{2.114}$$

The short-range interaction must vanish when electrons are far from each other

$$\lim_{r_{12}\to \infty} \hat{v}_{ee}^{SR}(r_{12}) = 0 \tag{2.115}$$

and, by assumption, at short-distance separations, the full interaction is recaptured by the operator

$$\lim_{r_{12}\to 0} \hat{v}_{ee}^{SR}(r_{12}) = \frac{1}{r_{12}}. \tag{2.116}$$

Under these constraints, the total electron interaction $\hat{V}_{ee}$ is decomposed accordingly as

$$\hat{V}_{ee} = \hat{V}_{ee}^{SR} + \hat{V}_{ee}^{LR}, \tag{2.117}$$

with

$$\hat{V}_{ee}^{SR/LR} = \sum_{i>j} \hat{v}_{ee}^{SR/LR}(r_{ij}). \tag{2.118}$$

This separation is intended to describe the long-range electron interaction using a MC wave function theory, while the short-range interaction is covered by DFT, leaving the electron-electron cusp region to be described by the DFA. The selection is supported by the short-range character of DC, correctly recovered by DFT, whereas local exchange density functionals exhibit unphysical asymptotic behavior at large interelectronic distances. On the other hand, as previously noted, the NDC caused by near-degeneracies of the ground and low-lying excited states, is correctly described by MC expansions [193].

The exact ground state range-separated density functional can be constructed by means of Levy's constrained search formalism [194]

$$E_0 = \min_{\rho} \left\{ \min_{\Psi\to\rho} \langle \Psi | \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{LR} | \Psi \rangle + E^{SR}[\rho] \right\}, \tag{2.119}$$

where $\hat{T}$ and $\hat{V}$ are the kinetic energy and electron-nuclei interaction operators and $E[\rho]^{SR}$ the short-range functional defined as the difference of universal functionals corresponding to full and long-range electron interaction operators [195, 78],

$$E[\rho]^{SR} = \min_{\Phi\to\rho} \langle\Psi|\hat{T} + \hat{V}_{ee}|\Psi\rangle - \min_{\Psi\to\rho} \langle\Psi|\hat{T} + \hat{V}_{ee}^{LR}|\Psi\rangle \,. \tag{2.120}$$

Double minimization in Eq. 2.119 is equivalent to minimization with respect to the antisymmetrized electronic wave function

$$E_0 = \min_{\Psi} \Big\{ \langle\Psi|\hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{LR}|\Psi\rangle + E^{SR}[\rho_\Psi] \Big\}, \tag{2.121}$$

with the electron density ρ_Ψ generated by the wave function Ψ. The wave function minimizing this expression (here denoted as Ψ^{LR}), is the ground state eigenfunction of the LR Hamiltonian

$$\hat{H}^{LR}\Psi^{LR} = \mathcal{E}_0\Psi^{LR}, \tag{2.122}$$

this Hamiltonian includes the long-range electron interaction operator and the short-range local potential $\hat{V}^{SR}$,

$$\hat{H}^{LR} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{LR} + \hat{V}^{SR}[\rho_{\Psi^{LR}}], \tag{2.123}$$

The local potential is computed from the functional derivative of the short-range density functional is given as

$$\hat{V}^{SR}[\rho] = \sum_i \frac{\delta E^{SR}[\rho]}{\delta\rho(\vec{\boldsymbol{r}}_i)}. \tag{2.124}$$

This wave function is free from the electron coalescence cusp (exact many-electron wave functions must exhibit cusps in regions where the coordinates of any two electrons coalesce [196]). The electron density $\rho_{\Psi^{LR}}$ generated by Ψ^{LR} coincides with the exact ground state density, ρ_0, as in the conventional DFT [197], $\rho_0 = \rho_{\Psi^{LR}}(\vec{\boldsymbol{r}})$. The total ground state energy obtained in the range separated formalism is given as the sum of the WFT and DFT parts,

$$E_0 = \left\langle\Psi^{LR}\middle|\hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{LR}\middle|\Psi^{LR}\right\rangle + E^{SR}[\rho_{\Psi^{LR}}]. \tag{2.125}$$

The short-range density functionals account for the short-range Hartree, $E_H^{SR}[\rho]$, and the exchange-correlation terms, $E_{xc}^{SR}[\rho]$ [198, 195],

$$E^{SR}[\rho] = E_H^{SR}[\rho] + E_{xc}^{SR}[\rho]. \tag{2.126}$$

For the implementation of this formalism, three crucial steps are of fundamental importance. These include the selection of a specific form of the long-range electron interaction operator, the development of a compatible short-range functional, and the selection of a WFT approximation to be employed in the variational equation 2.122. While multiple choices can be made for the splitting of the Coulomb operator [172, 199, 200, 201], one of the most comvenient splittings is based on the use of the error function (erf(x)) [173], being

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x du\ \exp\left(-u^2\right). \tag{2.127}$$

The Coulomb splitting can be written as

$$\frac{1}{r_{12}} = \frac{1 - \mathrm{erf}(\mu r_{12})}{r_{12}} + \frac{\mathrm{erf}(\mu r_{12})}{r_{12}}. \tag{2.128}$$

Within the use of the error function, the first term on the right-hand side (RHS) of this equation accounts for the short-range interaction, while the second term accounts for the long-range interaction in accordance with the conditions mentioned earlier. Within this choice, to optimize the separation between DC and ND correlations, it is crucial to choose the parameter μ optimally. Recent works [202, 203, 204] have provided recipes for determining an optimal μ for MCSCF wave functions in conjunction with short-range DFT (MC-srDFT). Initial results from range separation methods show promise, offering higher-quality results with computational costs scaling similarly to CASSCF [202, 203, 200, 201, 173, 171, 205, 206]. However, investigating chemical reactions or electronic excitations of transition metal complexes may face challenges due to steep scaling with the system size. A potential solution is combining srDFT with density matrix renormalization group approaches capable of handling larger active spaces in the lr-MC-WFT part [207]. In this thesis, we use the range separation formalism to combine the RASCI method with short-range DFT (srDFT). More details about our procedure can be consulted in subsection 4.3.1.

2.6.6 Adiabatic connection based methods

To finalize this review, it is worth mentioning recent approaches based on expressions derived from the adiabatic connection (AC). The AC formula for electron correlation energy provides a balanced treatment of the correlation energy, allowing for the accurate inclusion of DC effects in MR wave functions. The AC formalism, initially proposed in the KS-DFT framework (AC(KS-DFT)) [185, 208, 209], has been extended to cover single determinantal HF reference wave functions (AC(HF)) [210, 211] and more recently, to MC wave functions (AC(MC)) [212, 213]. The AC(MC) framework offers a computationally efficient alternative to second-order multireference perturbation methods, maintaining competitive accuracy [212, 213]. Unlike AC(KS-DFT) and AC(HF) methods limited to ground states of singlet spin symmetry, AC(MC) is applicable to both ground and excited states with arbitrary spin multiplicities [214, 70, 215]. The resulting computational cost of AC based methods makes them one of the most efficient approaches for accounting for a DC and NDC in a simultaneous balanced way [216].

In AC theory, a model Hamiltonian ($\hat{H}^{(0)}$) for which a reference wave function is one of its eigenfunctions, is assumed

$$\hat{H}^{(0)}\Psi^{ref} = E^{(0)}\Psi^{ref}. \tag{2.129}$$

Next, the parameter-dependent adiabatic Hamiltonian $\hat{H}^{\alpha}$ is constructed as

$$\hat{H}^{\alpha} = \hat{H}^{(0)} + \alpha\hat{H}'. \tag{2.130}$$

By varying the coupling strength parameter α between 0 and 1, one smoothly switches between the model and the exact Hamiltonians, respectively

$$\hat{H}^{\alpha=0} = \hat{H}^{(0)}, \tag{2.131}$$

$$\hat{H}^{\alpha=1} = \hat{H}. \tag{2.132}$$

Be $\{\Psi^{\alpha}_{\lambda}\}$ the set of eigenfunctions of the adiabatic Hamiltonian,

$$\hat{H}^{\alpha}\Psi^{\alpha}_{\lambda} = E^{\alpha}_{\lambda}\Psi^{\alpha}_{\lambda}, \tag{2.133}$$

Ψ^{α} denoting the particular wave function, which coincides with the reference wave function when $\alpha = 0$ ($\hat{H}^{\alpha}\Psi^{\alpha} = E^{\alpha}\Psi^{\alpha}$; $\Psi^{\alpha=0} = \Psi^{ref}$). Via the Hellman-Feynman theorem, it can be shown that the correlation energy follows from the following integration [217]

$$E_c^{AC} = \int_0^1 d\alpha\, W^{\alpha}, \tag{2.134}$$

where the adiabatic connection integrand W^{α} is

$$W^{\alpha} = \langle\Psi^{\alpha}|\hat{H}'|\Psi^{\alpha}\rangle - \left\langle\Psi^{ref}\middle|\hat{H}'\middle|\Psi^{ref}\right\rangle, \tag{2.135}$$

where W^{α} is the exact AC integrand. In the AC(KS-DFT) theory, the KS determinant is used as the reference wave function which delivers the exact electron density. The corresponding model Hamiltonian consists of the kinetic energy operator and the local KS potential. And the adiabatic connection Hamiltonian, in this case, includes the kinetic energy operator, a linearly scaled electron interaction operator, and a local α-dependent potential, which fixes the density to the exact full interacting density for each α

$$\hat{H}^{\alpha}_{KS-DFT} = \hat{T} + \alpha\hat{V}_{ee} + \hat{V}^{\alpha}_{loc} \qquad \text{such that} \quad \forall_{\alpha}\rho^{\alpha} = \rho_{exact}. \tag{2.136}$$

The AC(KS-DFT) integrand yielding the exact KS-DFT correlation energy takes the form

$$W^{\alpha}_{KS-DFT} = \langle\Psi^{\alpha}|\hat{V}_{ee}|\Psi^{\alpha}\rangle - E_{HX} \tag{2.137}$$

where E_{HX} is the sum of the Hartree and exchange energies.

The AC Hamiltonian in WFT is different from that in KS-DFT [218]. In general, WFT reference wave functions are multiconfigurational. In such cases, AC approaches aim at treating only the DC not accounted by the reference wave function. Thus, it has been justified to adopt another approximation. Particularly, when coupled with the extended random phase approximation (ERPA), the AC formalism allows for the computation of correlation energy using only the reference 1-RDM and 2-RDM. In this context, it is convenient to express the AC integrand in terms of transition RDMs $\gamma^{\alpha,\lambda}$, occupation numbers ν_p, and two-electron integrals $\langle rs|pq\rangle$ in the following form

$$W^{\alpha} = \frac{1}{2}\sum_{pqrs}{}' \left(\sum_{\lambda}\gamma^{\alpha,\lambda}_{pr}\gamma^{\alpha,\lambda}_{qs} + (\nu_p - 1)\nu_q\delta_{rq}\delta_{ps}\right)\langle rs|pq\rangle \tag{2.138}$$

where the prime indicates exclusion of terms involving all four indices p, q, r, s, belonging to a set of either active or inactive orbitals. The AC expression has been obtained under an assumption that the one-electron RDM stays constant along the AC path.

Extension to excited states can be performed as follows: if the reference state represents the Pth state, its connection with the vth state of the AC Hamiltonian is done in terms of the transition RDMs i.e., $\gamma^{\alpha,v}_{p,q} = \langle \Psi^{\alpha}_{P} | \hat{\psi}^{\dagger}_{q} \hat{\psi}_{p} | \Psi^{\alpha}_{v} \rangle$. In the approximate AC formalism, transition RDMs are obtained as [214]

$$(\gamma^{\alpha,v})_{p,q} = (\nu_p - \nu_q)[\vec{\boldsymbol{Y}}^{\alpha}_{v}]_{pq} \quad \forall_{p>q} \tag{2.139}$$

$$(\gamma^{\alpha,v})_{p,q} = (\nu_p - \nu_q)[\vec{\boldsymbol{X}}^{\alpha}_{v}]_{pq} \quad \forall_{q>p}, \tag{2.140}$$

where the vectors $\vec{\boldsymbol{Y}}^{\alpha}_{v}$ and $\vec{\boldsymbol{X}}^{\alpha}_{v}$ are solutions to the ERPA equations [219, 220] taking the form of a generalized eigen problem as

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \vec{\boldsymbol{X}}^{\alpha}_{v} \\ \vec{\boldsymbol{Y}}^{\alpha}_{v} \end{pmatrix} = \omega_v \begin{pmatrix} -\mathbf{N} & 0 \\ 0 & \mathbf{N} \end{pmatrix} \begin{pmatrix} \vec{\boldsymbol{X}}^{\alpha}_{v} \\ \vec{\boldsymbol{Y}}^{\alpha}_{v} \\ , \end{pmatrix} \tag{2.141}$$

The matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{N}$ are given for the α-dependent elements of the adiabatic connection Hamiltonian and 1- and 2-RDMs. Solutions of the ERPA equations come in pairs of positive and negative eigenvalues corresponding, respectively, to excitation and deexcitation energies [220].

The elements of $\mathbf{A}$ and $\mathbf{B}$ are linear in terms of the coupling strength parameter α,

$$\mathbf{A} = \mathbf{A}^{(0)} + \alpha \mathbf{A}^{(1)} \tag{2.142}$$

$$\mathbf{B} = \mathbf{B}^{(0)} + \alpha \mathbf{B}^{(1)}, \tag{2.143}$$

with $\mathbf{A}^{(0)}$, $\mathbf{A}^{(1)}$, $\mathbf{B}^{(0)}$, $\mathbf{B}^{(1)}$ being α independent. The overall matrices $\mathbf{A}$ and $\mathbf{B}$ are determined from the Hamiltonian and the reference wave function as follows

$$\mathbf{A}_{pq,rs} = \left\langle \Psi^{ref} \middle| [\hat{\psi}^{\dagger}_{p} \hat{\psi}_{q}, [\hat{H}, \hat{\psi}^{\dagger}_{s} \hat{\psi}_{r}]] \middle| \Psi^{ref} \right\rangle \quad \forall_{p>q\,,\,r>s} \tag{2.144}$$

$$\mathbf{B}_{pq,rs} = \left\langle \Psi^{ref} \middle| [\hat{\psi}^{\dagger}_{p} \hat{\psi}_{q}, [\hat{H}, \hat{\psi}^{\dagger}_{r} \hat{\psi}_{s}]] \middle| \Psi^{ref} \right\rangle \quad \forall_{p>q\,,\,r>s} \tag{2.145}$$

and the $\mathbf{N}$ matrix (metric matrix) is given by natural occupation numbers from the reference wave function,

$$\mathbf{N}_{pq,rs} = \delta_{pr} \delta_{qs} (\nu_q - \nu_p). \tag{2.146}$$

Using the ERPA approximated transition density matrices 2.139 and 2.140, the spin-summed AC integrand 2.138 (changing λ label in 2.138 by v) becomes

$$
\begin{aligned}
W^{\alpha} =& 2 \sum_{p>q,r>s}{}' \{(\nu_p - \nu_q)(\nu_r - \nu_s) \\
& \times \sum_{v} ([\vec{\boldsymbol{Y}}^{\alpha}_{v}]_{pq} - [\vec{\boldsymbol{X}}^{\alpha}_{v}]_{pq})([\vec{\boldsymbol{Y}}^{\alpha}_{v}]_{rs} - [\vec{\boldsymbol{X}}^{\alpha}_{v}]_{rs}) \\
& - \frac{1}{2}[\nu_p(1-\nu_q) + \nu_q(1-\nu_p)]\delta_{pr}\delta_{qs}\}\langle pr|qs\rangle.
\end{aligned}
\tag{2.147}
$$

This equation, in conjunction with the adiabatic connection formula for the correlation energy, provides a relatively straightforward approach for treating DC in a broad range of MC theories. It is important to note that the approach outlined here is not exact and relies on two approximations: (i) the assumption of the invariance of the one-electron density matrix along the AC path, and (ii) the utilization of transition density matrices from ERPA equations.

Considering that DC is anticipated to be described by the adiabatic connection, a further simplification can be made by approximating W^{α} to the lowest-order terms in the α expansion,

$$
W^{\alpha} = W^{(0)} + \alpha W^{(1)}, \tag{2.148}
$$

which results in the following expression for the correlation energy

$$
E_c^{AC0} = W^{(0)} + \frac{1}{2}W^{(1)}. \tag{2.149}
$$

For a general multireference wave function, these two terms can be found by applying perturbation theory to the ERPA eigenproblem. Another possible approximation for the AC integrand is based on the interpolation between the full-correlated limit and the correct no-intergroup-correlation (for which $W^{(0)} = 0$), i.e.,

$$
W^{\alpha} = \alpha W^{\alpha=1}, \tag{2.150}
$$

which leads to

$$
E_c^{AC1} = \frac{1}{2}W^{\alpha=1}. \tag{2.151}
$$

These two extrapolation schemes, AC1 and AC0, were recently proposed [216]. AC1 holds an advantage over the full AC approach, as the latter involves constructing ERPA matrices, solving ERPA equations, and computing integrals to determine correlation energy, all of which must be performed for multiple values of the coupling constant α. In contrast, AC1 requires these steps only at $\alpha = 1$. Furthermore, the AC formalism has recently been combined with the range separation of the Coulomb operator, resulting in a combined method that incorporates a novel long-range correction for DC effects. In this mingled approach, the short-range density functional part is introduced, with the corresponding DFA dependent on the on-top pair density [221].

2.7 Tools for analysis

The central focus of this thesis is on the development and implementation of innovative approaches for integrating DC treatment into hybrid schemes. To achieve this goal, it is essential to utilize available tools designed for analyzing electronic correlation and other relevant properties. In practice, the distinction between NDC and DC is often unclear; for this reason, numerous tools and diagnostics have been developed.

In the following subsections, we offer a comprehensive overview of selected tools and indices commonly used for analysis purposes. This detailed exploration aims to establish a reference of their functionalities and capabilities. This overview serves as an introductive framework, setting the stage for their integral role in the next chapters. As we engage in the discussion of our research, we will consistently utilize and refer to these tools.

2.7.1 Natural orbitals

Recall the expression of the 1-RDM (2.40) in terms of ladder operators,

$$\gamma(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}') = \langle \Psi | \hat{\psi}(\vec{\boldsymbol{x}}')^\dagger \hat{\psi}(\vec{\boldsymbol{x}}) | \Psi \rangle \,, \tag{2.152}$$

where $\vec{\boldsymbol{x}}$ refers to the space-spin coordinate of the electron. Among other properties of the 1-RDM, its diagonal elements, have the following physical interpretation: $\gamma(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}')d\vec{\boldsymbol{x}}_1$ = number of particles times the probability for finding a particle within the volume dv_1 around the point $\vec{\boldsymbol{r}}_1$ with spin s_1. If this density matrix is assumed to be the kernel of an integral operator $\hat{D}$, then

$$\hat{D} f(\vec{\boldsymbol{x}}) = \int d\vec{\boldsymbol{x}}' \gamma(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}') \, f(\vec{\boldsymbol{x}}'). \tag{2.153}$$

From this point of view, as an integral kernel, the 1-RDM is a bounded linear Hermitian operator with an infinite but countable eigenvalue spectrum. $\hat{D}$ is thus Hermitian and nonegative, whose rank (number of nonzero eigenvalues) may be finite or infinite. The eigenfunctions $\{\chi_p\}$ of $\hat{D}$, defined by [36, 222, 223]

$$\int dx' \gamma(\vec{\boldsymbol{x}}, \vec{\boldsymbol{x}}') \chi_q(\vec{\boldsymbol{x}}') = \hat{D}\chi_q(\vec{\boldsymbol{x}}) = \nu_q \chi_q(\vec{\boldsymbol{x}}), \tag{2.154}$$

are called the natural spin orbitals (NSOs) of Ψ. The respective eigenvalues are the natural spin orbital occupation numbers (NSOONs). Once diagonalized, the 1-RDM can be written in the form

$$\gamma(\vec{\boldsymbol{x}}_1', \vec{\boldsymbol{x}}_1) = \sum_q n_q \chi_q^*(\vec{\boldsymbol{x}}_1') \chi_q(\vec{\boldsymbol{x}}_1). \tag{2.155}$$

NSOONs represent the average number of particles in each one of the natural spin orbitals. Occupation numbers fulfill two conditions:

$$0 \leq \nu_p \leq 1, \sum_p \nu_p = N. \tag{2.156}$$

Consequently, the total number of particles (N) can be distributed among more than N spin orbitals. A limiting case is recognized when exactly N spin orbitals are occupied, this occurs under the conditions

$$\gamma^2 = \gamma,\ \mathrm{Tr}[\gamma] = N. \tag{2.157}$$

The former is the idempotency condition of the 1-RDM, while the latter refers to the fact that, the sum of the diagonal elements must sum to the total number of particles in the system (N). When these two conditions are fulfilled, the sum of the diagonal elements has exactly N eigenvalues equal to 1 and the remaining are zero. In fact, idempotency is a condition satisfied by pure states represented by a single Slater determinant [124]. If we choose the NSOs as the basis set, and if conditions 2.157 are satisfied, then the CI expansion (equation 2.27) is reduced to a single determinant. The fact that occupation numbers can display non-integer values between 0 and 1 is a sign that multiple configurations also contribute to the CI expansion. This is one of the most distinct features of interacting systems. The presence of fractional natural occupation numbers reflects the electronic correlation present in the system under consideration. If the system is weakly correlated, occupation numbers deviate slightly from 0 or 1, while for strongly correlated systems, large deviations from such integers are expected.

Consider a wave function Ψ that is an eigenfunction of $\hat{S}^2$ and $\hat{S}_z$, with eigenvalues $S(S+1)$ and M. The spatial dependence of the corresponding NSOs is computed from different integral kernels for α and β spins. Only if $M = 0$ will the space parts of the α NSOs coincide with the space part of the β NSOs, and they will differ for other values of M.

For CI calculations, it is preferable to use the spin-free counterpart, the natural orbitals (NO), defined as the eigenfunctions of the spinless 1-RDM $\gamma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r}}')$.

$$\int dr' \gamma(\vec{\boldsymbol{r}}, \vec{\boldsymbol{r'}})\phi_p(\vec{\boldsymbol{r'}}) = \nu_p \phi_p(\vec{\boldsymbol{r}}), \tag{2.158}$$

the eigenvalues ν_p, are those now between 0 and 2, represent the natural orbital occupation numbers (NOONs). One of the primary advantages of employing NOs in CI calculations stems from the expectation that CI expansions based on these orbitals exhibit the fastest convergence. In other words, for a given level of accuracy, NOs yield the most compact expansion possible. This characteristic was initially demonstrated by Löwdin [36] in the context of a two-electron system, where the basis of natural orbitals requires the fewest configurations to achieve a specified level of accuracy in the energy.

While achieving such improvements may be challenging in a general multielectron case (three or more electrons) due to a considerable reduction in the number of determinants only occurring for specific restricted cases [222, 224], cumulative evidence from previous studies indicates a notably faster convergence when utilizing NOs (for example, see [225, 226, 227, 228, 229]).

The properties of NOs make them a desirable basis for various purposes. Particularly, it has been shown [224] that they provide an effective partitioning of the virtual orbital space into subspaces suitable for correlating different occupied shells. This allows truncation of the set of orbitals used in the CI expansion with a considerably smaller effect on the result than with a truncation in terms of canonical SCF orbitals [225, 229].

In the context of active space selection in quantum chemical calculations, a notable strategy involves using NOONs as a guide. Specifically, NOONs can inform about the choice of an active space, as demonstrated in CAS theories [230, 156]. Notably, Pulay and colleagues [156] propose the straightforward and computationally efficient approach of employing NOs obtained from unrestricted Hartree-Fock (UHF) calculations. Assuming successful convergence, a practical implementation involves selecting those orbitals with fractional occupancy, typically ranging between 1.98 and 0.02. These selected orbitals are then employed to define the active space in subsequent MCSCF calculations. However, it is important to note that achieving convergence in UHF solutions (a local minimum with respect to orbital rotations) for general systems can pose challenges, such as those encountered near the Coulson-Fisher point [231, 130], where the HF instabilities can be related to system's attributes like the geometry and the number π-electrons (in some cases) [232].

Extending this strategy, a similar methodology has been introduced for the determination of the active space in the RASSCF method. In this approach, the starting point involves an SCF density matrix that is averaged over the electronic states of interest. Subsequently, the choice of a RAS partition is formed by considering the corresponding natural orbital occupation numbers and applying Pulay's criterion [233].

Among various applications, NOONs can serve as a diagnostic tool for wave functions, especially in cases requiring multireference treatment. For fully variational wave functions (for example, RHF, MCSCF or FCI cases), satisfying the Hellmann-Feynman theorem, the first-order density is uniquely defined. In the case of nonvariational methods, not satisfying the Hellmann-Feynman theorem, there are various definitions of the first-order density, coming from the expected value of the particle number operator and from the derivative expression for the relaxed-density. Such cases can lead to nonphysical NOONS [234], mainly characterized by their ill-behavior along potential energy surfaces and by values larger than 2.0, or negative.

Results taken from model systems computed with single reference methods demonstrate the use of nonphysical NOON (mainly, negative) as an indicator of the breakdown of the single configuration approximation [234, 129]. Moreover, it has been observed that While improving the basis set does result in slight quantitative modifications, the problems associated to the presence of unphysical NOONs are unaltered [129].

Additionally, NOONs find utility in defining correlation indices [36, 235, 236, 46, 237, 238] and metrics for quantifying the amount of effectively unpaired electrons [239, 240, 241].

As we will illustrate in subsequent subsections, the natural orbital concept has expanded to a similar class of orbitals derived from the singular value decomposition of one-particle transition density matrices. These orbitals prove to be well-suited for describing excited states, playing a pivotal role in the analysis of such states.

2.7.2 Correlation indices based on natural orbitals

The quest for an efficient correlation metric is a current challenge in the electronic structure community. Various statistical measures, including leading CI coefficients (c_i), CC amplitudes (t_i^a, t_{ij}^{ab}), and natural orbital occupations (ν_i), have been proposed. While much effort has focused on establishing metrics for NDC [242, 243, 244, 245, 246, 247, 156, 155, 248], there are fewer works on DC [45, 46, 249, 237, 250].

Among these criteria, the use of natural orbitals and their occupations is a desired choice due to their straightforward determination (involving only the diagonalization of the 1-RDM)

and interpretation (providing qualitative insights and spatial localization about correlated electrons).

In this thesis, we adopt numerical indices based on the use of NOONs, as recently proposed in our group [46, 237], to account for DC and NDC, denoted respectively as I_D and I_{ND}. The concept raises from the ability to decompose electron correlation into dynamic and nondynamic components using the cumulant of the 2-RDM. Under this scheme, global I_D and I_{ND} indices can be defined from a two-electron model [46],

$$I_D = \frac{1}{4}\sum_{\sigma,p}[\nu_p^\sigma(1-\nu_p^\sigma)]^{1/2} - \frac{1}{2}\sum_{\sigma,p}\nu_p^\sigma(1-\nu_p^\sigma) \tag{2.159}$$

$$I_{ND} = \frac{1}{2}\sum_{\sigma,p}\nu_p^\sigma(1-\nu_p^\sigma), \tag{2.160}$$

such that

$$I_{total} = I_D + I_{ND} = \frac{1}{4}\sum_{\sigma,p}[\nu_p^\sigma(1-\nu_p^\sigma)]^{1/2}, \tag{2.161}$$

where ν_p^σ is the occupation of the p-th NSO with spin σ. Notice that I_{ND} serve as a measure of the deviation from the idempotency of the 1-RDM (2.40), indicating the lack of a single-reference representation of the wave function. These expressions assume that for an N-electron system, the total measure can be obtained by summing individual orbital contributions.

Previously, it has been shown one important feature of I_D and I_{ND}, their size-extensivity [251], since they are able to display correct asymptotic behavior in the study of dimers (i.e. at infinite separations, their values are exactly the sum of the respective monomer values). However, I_D and I_{ND} are not size intensive, and their values tend to scale with system size.

The simplicity of these electron correlation functions makes them applicable to a wide range of electronic structure methods, provided a set of natural orbitals and occupancies are available. A normalized version of these indices (denoted by $\bar{I}_D$ and $\bar{I}_{ND}$) has been proposed in our group (see Ref. [252]). These will be used in Chapter 6 as an auxiliary tool in the analysis of a method combining CASSCF wave functions with an effective potential for DC.

2.7.3 Fractional occupation density

In practice, the analysis of strongly correlated electrons characterized by a multiconfigurational wave function can be challenging due to the unclear distinction between NDC and DC contributions to a system. As noted previously, NOs and their NOONs might play an essential role in measuring electronic correlation. Occupation values provide information about static and dynamic correlation through the number of unpaired electrons, while associated NOs can offer insight into the spatial distribution of correlated electrons.

For chemical purposes, elucidating the number of unpaired electrons and their spatial distribution is of interest for related properties such as diradical character or chemical stability. The fractional occupation number weighted density (FOD), introduced by Hansen and Grimme [253, 254], is a relatively simple tool designed for the real-space diagnostic of NDC. The FOD can be generaly described by

$$\rho^{FOD}(\vec{\boldsymbol{r}}) = \sum_{p} f(\nu_p)\phi_p^*(\vec{\boldsymbol{r}})\phi_p(\vec{\boldsymbol{r}}), \tag{2.162}$$

where, $f(\nu_p)$ represents the fractional electron distribution number, a function of the electronic occupancy of the p-th natural orbital. The metric for quantifying unpaired electrons relies on the specific ν_p to $f(\nu_p)$ mapping chosen. Widely employed forms of $f(\nu_p)$ are those defined by Yamaguchi (f_Y) [239]

$$f_Y(\nu_p) = 2\nu_p - \nu_p^2 \tag{2.163}$$

and the extensions proposed by Head-Gordon (f_U and f_S) [240]

$$f_U(\nu_p) = \min(\nu_p, 2 - \nu_p), \tag{2.164}$$

$$f_S(\nu_p) = \nu_p^2(2 - \nu_p)^2. \tag{2.165}$$

The effective number of unpaired electrons in electronic states can be determined by examining the NOONs ($0 \leq \nu_p \leq 2$) using the expression

$$N = \sum_{p} f(\nu_p) = \int d\vec{\boldsymbol{r}}\rho^{FOD}(\vec{\boldsymbol{r}}), \tag{2.166}$$

where the sum runs over all natural orbitals and the contributions of the doubly occupied and unoccupied orbitals are exactly zero [255]. f_Y and f_U metrics can be inserted into equation 2.166 to compute the total number of unpaired electrons, denoted as N_Y and N_U, respectively. Moreover, they will produce the same values for $\nu_p = 0$, 1, 2, but will differ at intermediary values. It has been argued that f_Y mapping tends to overestimate the number of unpaired electrons, while f_U should be more capable to provide reliable results [240, 255, 115].

2.7.4 Natural transition orbitals

In the study of excited states, the interaction among electrons introduces a layer of complexity when attempting to understand electronic transitions. As electrons interact with each other, the alteration of one electron's state affects the states of the remaining electrons. This complexity poses challenges to traditional descriptions of electronic transitions, such as the inaccuracies found in Koopmans' model. Koopmans' model envisions electronic excitations as the promotion of an electron from occupied orbitals to virtual ones. However, this model is not only quantitatively inadequate, but also qualitatively insufficient in many scenarios [256].

An additional challenge can be illustrated when considering the lowest level of theory capturing the physics of excited states, the CIS wave function, constructed as the linear combination of singly excited Slater determinants Ψ_i^a, computed from a reference HF ground-state determinant,

$$\Psi^{CIS} = \sum_{i,a} C_i^a \Psi_i^a. \tag{2.167}$$

The analysis of the leading amplitudes, C_i^a, in this expansion facilitates the interpretation of excited states in terms of hole and particle orbital transitions. For example, a CIS/6-31G* calculation of ethylene shows that the lowest excited state can be described as a $\pi \rightarrow \pi^*$ transition (the weight of the respective $\Psi_i^a \sim 0.94$) [256].

Unfortunately, analyses based on expansion amplitudes, face challenges in terms of interpretative power. While the CIS energy and excited-state properties are constant under unitary transformations of orbitals within occupied or virtual spaces, the wave function amplitudes prove to be sensitive to orbital choices, giving rise to a basis-set dependence. As in the example of ethylene, when the basis is extended to 6-31+G*, several Rydberg orbitals appear below the π^* valence MO and the CIS wave function of the lowest excited triplet state has two comparable amplitudes involving both Rydberg and valence orbitals.

Since the wave function amplitudes are not observables. A more robust analysis can be achieved by considering experimental observables such as the transition density ($\rho(\vec{\boldsymbol{x}}_e, \vec{\boldsymbol{x}}_p)$) among an initial and a final state,

$$\rho(\vec{\boldsymbol{x}}_e, \vec{\boldsymbol{x}}_h) = \sum_{pq} \gamma_{pq}^{FI} \phi_p(\vec{\boldsymbol{x}}_e) \phi_q(\vec{\boldsymbol{x}}_h), \tag{2.168}$$

where $\vec{\boldsymbol{x}}_e$ and $\vec{\boldsymbol{x}}_h$ denote the electron (particle) and hole coordinates, γ_{pq}^{FI} is a one-particle transition density matrix (1PTDM) defined as

$$\gamma_{pq}^{FI} = \langle \Psi_F | \hat{\psi}_p{}^{\dagger} \hat{\psi}_q | \Psi_I \rangle \tag{2.169}$$

where I labels initial state and F the final one, σ as before, is the spin index. Here $\hat{\psi}_{p\sigma}^{\dagger}$ and $\hat{\psi}_{q\sigma}$ are creation and annihilation operators corresponding to the ϕ_p and ϕ_q molecular orbitals, respectively. The transition density is related to observable properties because it defines the probability of transition. In the case of the 1PTDM, this is a rectangular non-Hermitian matrix whose most compact representation can be achieved (in analogy to the diagonalization process to obtain the NOs from the 1PDM), through singular value decomposition (SVD) [257] by

$$\gamma = \mathbf{V} \mathbf{\Sigma} \mathbf{U}^{\mathbf{T}}, \tag{2.170}$$

where $\mathbf{\Sigma}$ is the diagonal matrix of singular values, λ_q. $\mathbf{V}$ and $\mathbf{U}$ contain the hole ($\psi_K^h(\vec{\boldsymbol{x}})$) and particle (electron, $\psi_K^e(\vec{\boldsymbol{x}})$) natural transition orbitals (NTOs), according to

$$\psi_K^h(\vec{\boldsymbol{x}}) = \sum_q U_{qK} \phi_q(\vec{\boldsymbol{x}}) \tag{2.171}$$

and

$$\psi_K^e(\vec{\boldsymbol{x}}) = \sum_q V_{qK} \phi_q(\vec{\boldsymbol{x}}). \tag{2.172}$$

NTOs offer a means to describe electronic excitations in terms of molecular orbitals for any type of wave function [258]. Each pair is characterized by a hole in the occupied space associated with a corresponding single particle in the virtual space. Through this transformation, the transition density (Eq. 2.168) is reduced to

$$\rho(\vec{\boldsymbol{x}}_e, \vec{\boldsymbol{x}}_h) = \sum_K \lambda_K \psi_K^e(\vec{\boldsymbol{x}}_e)\psi_K^h(\vec{\boldsymbol{x}}_h), \tag{2.173}$$

the set $\{\lambda_q\}$ can be interpreted as the weights of the respective NTO pair when divided by the square of the Frobenius norm (Ω) of γ defined as

$$\Omega = ||\gamma|| = \sum_{pq} \gamma_{pq}^2 = \sum_K \lambda_K^2, \tag{2.174}$$

In fact, for a given excitation, the number of NTOs with significant weights reveals its inherent collectivity, representing the quantity of excitations characterizing a particular excited state. Among other indices, participation ratios (PR_{NTO}), have been introduced to account for this collectivity [256],

$$PR_{NTO} = \frac{(\sum_K \lambda_K^2)^2}{\sum_K \lambda_K^4} = \frac{\Omega^2}{\sum_K \lambda_K^4}. \tag{2.175}$$

For a transition involving a CIS final state with exactly one configuration, $PR_{NTO} = 1$ (pure excitations). If the CIS final state is described by two configurations, $PR_{NTO} = 2$, and so on. While pure excitations are characterized by $PR_{NTO} \approx 1$, transitions of mixed character will display values greater than unity. In Chapter 7, NTOs will serve as analytical tools used to identify the leading electronic transitions involved in cases where spin-orbit coupling (SOC) is enhanced or suppressed in selected diradical systems.

2.7.5 Coulson's electron correlation hole

The correlation in the movement of electrons means that the presence of one electron will impact the likelihood of finding another electron nearby. In this sense, the mere existence of an electron results in a decrease in the probability of locating another electron in its vicinity. In other words, an electron creates an area of reduced probability in its surroundings, often referred to as a "hole". This phenomenon is attributed to the Pauli principle and the fact that electrons are charged particles, giving rise to repulsive Coulombic and exchange interactions [259, 260, 40, 17]. A way to describe this is through the use of the pair density, which considers the simultaneous distribution of the electrons. For an N-electron system, described by the wave function $\Psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_N)$ the pair density is

$$\rho_2(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2) = N(N-1) \int d\,\vec{\boldsymbol{x}}_3 \ldots d\,\vec{\boldsymbol{x}}_N |\Psi(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2, \ldots, \vec{\boldsymbol{x}}_N)|^2, \tag{2.176}$$

where factor $N(N-1)$ corresponds to the McWeeny's normalization [261], accounting for the number of electron pairs in the system. The pair density retrieves information about the probability of finding electron 1 at the position and spin $\vec{\boldsymbol{x}}_1$ and electron 2 at the position and spin $\vec{\boldsymbol{x}}_2$ when the positions of the remaining $N-2$ electrons are averaged over the whole space. For Hamiltonians that contain at most two body interactions, the pair density determines all the relevant physical properties of the system. The number of variables (six spatial and two spin variables) in the pair density makes its analysis a complicated task. In order to use the pair density to define electron-pair distribution functions, we consider two electrons placed at $\vec{\boldsymbol{r}}_1$ and $\vec{\boldsymbol{r}}_2$. The interaction between them can be analyzed (conveniently) in terms of the center of coordinates and the distance between the two electrons, respectively

$$\vec{\boldsymbol{R}} = \frac{1}{2}(\vec{\boldsymbol{r}}_1 + \vec{\boldsymbol{r}}_2) \tag{2.177}$$

and

$$\vec{\boldsymbol{s}} = \vec{\boldsymbol{r}}_2 - \vec{\boldsymbol{r}}_1, \tag{2.178}$$

$\vec{\boldsymbol{R}}$ and $\vec{\boldsymbol{s}}$ are typically referred as extracule and intracule coordinates, respectively. With these variables, the dimensionality of the pair density can be reduced through space integration in the form

$$E(\rho_2, \vec{\boldsymbol{R}}) = \int d\vec{\boldsymbol{r}}_1 d\vec{\boldsymbol{r}}_2 \rho_2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)\delta(\vec{\boldsymbol{R}} - ((\vec{\boldsymbol{r}}_1 + \vec{\boldsymbol{r}}_2)/2)) \tag{2.179}$$

and

$$I(\rho_2, \vec{\boldsymbol{s}}) = \int d\vec{\boldsymbol{r}}_1 d\vec{\boldsymbol{r}}_2 \rho_2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)\delta(\vec{\boldsymbol{s}} - (\vec{\boldsymbol{r}}_2 - \vec{\boldsymbol{r}}_1)), \tag{2.180}$$

where δ is the Dirac delta function. Integration of the form 2.179 defines the so-called extracule probability density (EDP), while the integration of the form 2.180 defines the intracule probability density (IPD). Since EPD is a function of $\vec{\boldsymbol{R}}$, its evaluation provides information about the location of electron pairs. Particularly $E(\vec{\boldsymbol{R}})|_{R=0}$ is the so-called interelectronic counterbalance density, and has been proposed as an indicator of NDC [262].

Further integration of $I(\rho_2, \vec{\boldsymbol{s}})$ over the solid angle $d\Omega_u = \sin\theta_u d\theta_u d\phi_u$ (with u being $\vec{\boldsymbol{s}}$), defines what is called the isotropic radial intracule,

$$I(\rho_2, s) = s^2 \int d\Omega_u I(\rho_2, \vec{\boldsymbol{s}}). \tag{2.181}$$

This is a reduced form of the pair density that only depends on the interelectronic distance, s and corresponds to the probability distribution of interelectronic distances between the two electrons. As will be noted in the following, electron-pair distributions are useful tools for the description of electronic structure and electron correlation. In analogy to the electron correlation energy, Coulson defined the correlated pair density as the difference between the exact and the HF pair densities [263],

$$\Delta\rho_2^c(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2) = \rho_2(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2) - \rho_2^{HF}(\vec{\boldsymbol{x}}_1, \vec{\boldsymbol{x}}_2) \tag{2.182}$$

The intracule of the correlated pair is known as the Coulson's Coulomb hole [264, 265, 259]

$$h_c(s) = I(\Delta\rho_2^c, s) = I(\rho_2, s) - I(\rho_2^{HF}, s) \tag{2.183}$$

For a given basis set, the difference of the FCI intracule and the HF one pictures the lack of correlation in the HF approach and, since HF includes exchange interactions, this correlation hole is known as the Coulomb hole. Due to normalization, the integration of the Coulomb hole is zero ($\int_0^\infty ds\, h_c(s) = 0$). This property illustrates the usefulness of Coulomb holes as tools for analysis, since it implies that there is a fixed amount of electrons that are shifted from short- to long-range. Coulomb holes, in essence, offer a visual representation of the

probability density difference. This distinction serves as an indicator of the impact on the pair density resulting from the transition from the mean-field HF approximation, which underestimates electronic repulsion, to a correlated framework. Thus, as depicted in Fig. , qualitative comparisons among different methods can be stated in terms of Coulomb holes, providing valuable insights into the complexities of the electronic distribution obtained from a given approach.

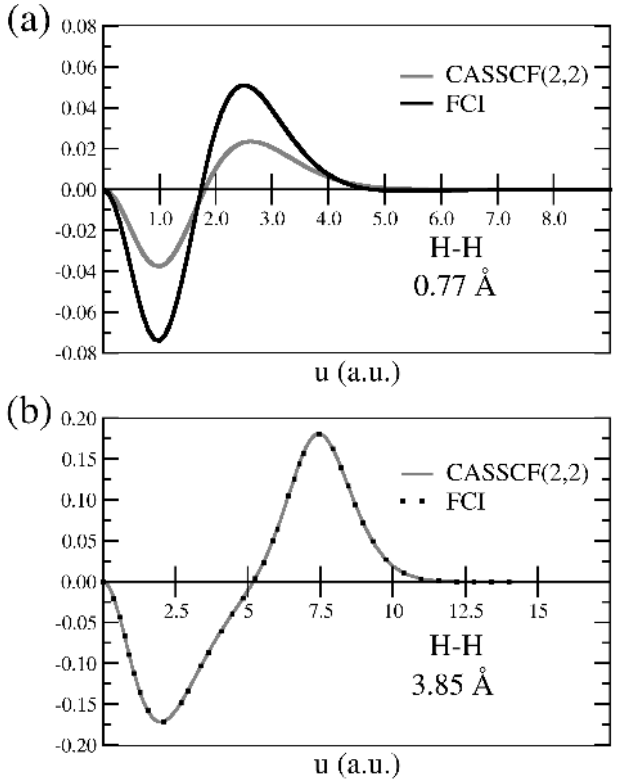


Figure 2.2: (a) Ilustrative comparison of CASSCF and FCI Coulomb holes computed for the H_2 at equilibrium distance (0.77 Å), the lack of dynamic correlation in the CASSCF wave function produces a qualitative change in the shape of the hole with respect to the FCI solution. (b) Same comparison computed at large separation (3.85 Å), in this case both solutions retrieve the same hole. Calculations performed with the cc-pVDZ basis set.

In chapter 6 of this thesis, we compute the intracules and corresponding Coulomb holes, in order to analyse and establish the effect of an effective potential to the electronic distribution of systems of interest.

As seen so far, the characterization of correlation energy present in a given state can be explored from different perspectives. While Coulomb holes serve as powerful tools for qualitative elucidation [266], the intricacies involved in their computation might present some difficulties in their determination or use (i.e., the need for pair densities in the computation of intracules).

CHAPTER 3

Objectives

This thesis represents an extension of previous works [267, 22, 23, 24], proposing novel methods for the simultaneous treatment of nondynamic and dynamic correlation within a unified approach that incorporates different methodologies. Furthermore, the scope of this thesis includes an expanded application of a recent methodology for computing SOCs [268].

The primary motivation for this research is to develop cost-effective electronic structure methods for the accurate and efficient computational study of ground and excited states. On one hand, our aim is to develop mingled WFT+DFT methods capable of effectively accounting for dynamic and nondynamic correlation while avoiding the double counting of electron correlations. On the other hand, guided by the performance achieved in previous approaches [26, 27, 28], we focus on implementing and developing a method based on the inclusion of dynamic correlation through an effective potential that directly links electron correlation to the wave function.

Concerning mingled WFT+DFT methods, we are expanding upon a previous idea based on the range separation of the Coulomb operator to merge WFT with DFT, referred to as WFT-*sr*DFT. While this method has been designed to effectively account for dynamic correlation at short-range interelectronic distances using a range separation parameter, alternative formulations intended to introduce spin polarization effects have been largely overlooked. Therefore, our first goal is to identify the problems and limitations of WFT-*sr*DFT in introducing these effects and to emphasize their importance in providing a balanced description between closed and open-shell configurations. To achieve this, we introduce different strategies based on the spin polarization of the electron density used in density functionals, and the re-definition of short-range exchange. And we aim to validate them by assessing their performance in the study of model open-shell systems.

Secondly, considering the advancements in WFT-*sr*DFT, we introduce an alternative WFT+DFT formulation that involves partitioning the natural orbital (NO) space. Large and small natural orbital occupation numbers (NOONs) are linked with WFT and DFT, respectively, with RASCI chosen as the WFT approach. In this context, our objective is to design, implement, and test various partitions of the NO space, while ensuring a suitable choice of the correlation functional.

Concerning the use of an effective potential for treating the dynamic correlation, our first goal is to reproduce the previous idea initially stated by Roos [24]. This approach requires

the evaluation of modified Gaussian damped electron repulsion integrals. We then intent to establish a theoretical relationship between the parameters within the effective potential and the effects it induces. Additionally, we aim to articulate prior knowledge concerning its potential impacts on energy and electronic structure. Our next goal will be to devise a parameterization scheme that yields optimal parameters, ensuring computed energies reach FCI quality. We extend Roos' work by introducing a further separation of the effective potential to discriminate in between inter- and intra-atomic electron pair contributions. Hence, an additional objective is to discern the significance and repercussions of this separation, achieved through additional parametrization efforts that provide a spectrum of optimal parameters for ensuring accurate results of FCI quality.

Additionally, the significance of SOC in various photophysical and photochemical processes, along with the performance achieved in describing open-shell systems, has motivated the implementation of novel and efficient approaches based on the spin-orbit mean field (SOMF) approximation. Our first goal in this regard is to expand the application of this strategy to the study of SOC in open-shell systems displaying di- or polyradical character. To this end, calculations were performed within non-relativistic wave functions obtained with the RASCI method [268]. We pursuit to unravel the origin of SOCs in complex scenarios within the help of additional analysis tools, such as the natural transition orbitals from the spinless transition density matrix.

In summary, this thesis addresses these questions by developing, implementing, testing, and applying methods to different atomic and molecular systems where the proper treatment of static and dynamic correlation becomes crucial. The objectives include:

- Extend the understanding and identify the effects of including spin polarization effects into WFT+DFT methods based on the range-separation of the Coulomb operator. And identify problems and limitations, emphasizing their significance in achieving a balanced description of closed and open-shell states.
- Develop, implement, and test an alternative WFT+DFT formulation that involves partitioning the NO space, through association of large and small NOONs with WFT and DFT, respectively.
- Develop and implement a method based on the use of an effective potential capable of treating dynamic electron correlation. Establish a theoretical relationship between effective potential parameters and their effects on energy and electronic structure.
- Testing the new methodologies on model systems to identify possible improvements, parametrization schemes and potential challenges.
- Expand the application of the SOMF approximation to study the origin of SOCs in open-shell systems with di- or polyradical character.

CHAPTER 4

Short-range DFT energy correction to multiconfigurational wavefunctions for open-shell systems

4.1 Abstract

Electronic structure methods emerging from the combination of multiconfigurational wave functions and density functional theory (DFT) aim to take advantage of the strengths of the two nearly antagonistic theories. One of the common strategies employed to merge wave function theory (WFT) with DFT relies on the range-separation of the Coulomb operator, in which DFT functionals take care of the short-distance part, while long-range inter-electronic interactions are evaluated by the chosen wave function method (WFT-*sr*DFT). In this work, we uncover the limitations of WFT-*sr*DFT in the characterization of open-shell systems. We show that spin polarization effects have a major impact on the (short-range) DFT exchange energy and are of vital importance in order to provide a balanced description between closed and open-shell configurations. We introduce different strategies to account for spin polarization in the short-range based on the definition of a spin polarized electron density and with the use of short-range exact exchange. We test the performance of these approaches in the dissociation of hydrogen molecule, the calculation of energy gaps in spin-triplet atoms and molecular diradicals, and in the characterization of low-lying states of the gallium dimer. Our results indicate that the use of short-range DFT correlation in combination with a (full-range) multiconfigurational wave function might be an excellent approach for the study of open-shell molecules, and largely improves the performance of WFT and WFT-*sr*DFT.

4.2 Introduction

Molecular systems with open-shell structures, *i.e.*, featuring unpaired electrons, are of major relevance in chemistry as they present unique electronic [269, 270] and magnetic properties [271], and play a major role as intermediates in chemical reactions as transition states or reaction intermediates [272, 273]. Many transition metal compounds exhibit partially

filled electronic levels which allows them to form paramagnetic compounds with many different applications such as in quantum computation and high-density data storage [274]. Unpaired electrons are also of major interest in organic chemistry, *e.g.*, diradicals (two unpaired electrons [275]) and polyradicals (multiple unpaired electrons), as they are strongly related to chemical stability and trigger a rich diverse reactivity [276]. Lately, there have been a renewed interest for the design of organic diradical and polyradicals because of the many potential applications of these compounds in various fields [277] such as molecular electronics [278, 279] and optoelectronics [280], nonlinear optics [269], magnetism [281], singlet fission [282, 283, 284] and the activation of molecular hydrogen [285, 286].

The electronic structure of open-shell systems is rather complicated compared to the close-shell form of many stable molecules, and pose an arduous challenge for quantum chemistry methods. Diradicals and polyradicals hold degenerated or near-degenerated orbitals at the occupied-virtual interphase of the orbital space, resulting in small energy gaps between several electronic states with different spin multiplicities. As a consequence, the electronic structure of these systems is of multiconfigurational nature and cannot be described by standard single-reference methods. Single determinant wave functions like the mean-field HF model are unable to characterize low-spin states of open-shell systems. Moreover, single-reference methods emerging from HF, *e.g.*, Møller–Plessed (MP) perturbation theory or truncated CI and CC expansions, cannot recover, in general, strong electron correlation effects between unpaired electrons. The multiconfigurational nature of diradicals and polyradicals can be described by WFT approaches built from two or more Slater determinants (multireference methods [287, 58]), like in the MCSCF family, with the CASSCF method [58, 288, 289] probably being its most ubiquitous example. These methodologies are very convenient for acquiring static correlation, but dynamic correlation converges very slowly with respect to the number of configurations. MCSCF solutions can be refined by including missing dynamic electron correlation contributions through CI schemes, like in MRCI [55, 56], or via multi-reference perturbation theory [290, 291, 59], like in CASPT2 [59, 292, 149, 150]. But the application of these methods is in many cases limited by the severe computational demands, preventing their use in medium to large molecules.

DFT [9, 293], and in particular its most popular KS formulation [9], presents an advantageous scaling of the computational requirements with respect to correlated WFTs and in many cases the accuracy of DFT-based results is comparable to high hierarchy WFT methods. For these reasons, KS-DFT has become the method of choice for large and complex molecular systems. In practice, DFT approximations are known for their ability to account for dynamic correlations, but show strong limitations to deal with degeneracies at the Fermi level linked to strong (non-dynamic) correlation [160], which are essential to characterize open-shell systems. The different properties of the two seemingly orthogonal theories have motivated many researchers to explore the possibilities of WFT-DFT mingled methodologies [294].

One of the possible strategies for including DFT correlation effects into electronic wave functions from WFT is based on the range separation of the Coulomb electron-electron interaction operator into short and long inter-electron distance contributions [200]. The use of range-separation to combine DFT with MCSCF relies on the fact that dynamic correlation effects can be related to instantaneous electron correlations at short inter-electronic distances, which are expected to be well described by DFT energy functionals, while the long-range part is better evaluated by WFT. This methodology, often referred as WFT-*sr*DFT, has been developed and applied rather successfully in numerous occasions [295, 296, 208, 297, 298, 78, 173, 76, 79, 203, 267, 299, 300, 301, 302, 303, 153].

In the present study, we explore in great detail the performance of WFT-*sr*DFT in the computation of relative energies in open-shell systems. Identifying the general problems and limitations of this approach enable us to propose several alternative formulations within the WFT-*sr*DFT frame and study their validity in several paradigmatic atomic and molecular systems with strong diradical character. To that aim, here we use a RASCI wave function [304, 125, 126, 151] at the long range combined with short-range exchange and correlation functionals.

This Chapter is organized as follows. Section 4.3.1 discuses the theoretical foundations of WFT-*sr*DFT, presents a state average approximation for the calculation of several eigenstates, exposes some of the problems and their physical origin of WFT-*sr*DFT for open-shell systems using H_2 as a toy model, and introduces potential alternatives to overcome these limitations based on polarization of the spin density and the re-definition of short-range exchange. Section 4.4 describes the technical details of the performed calculations. Then, we present and discuss the results obtained with WFT and the various WFT-*sr*DFT schemes in the study of H_2 dissociation, energy gaps in atoms, small diradicals, benzyne isomers and trimethylenemethane molecule, and the low-lying states in the gallium dimer. Finally, in section 4.6 we state the main conclusions of the chapter.

4.3 Theory

4.3.1 Range separation in WFT-*sr*DFT

An elegant strategy to combine multiconfigurational wave functions with DFT is to employ range separation. This strategy is based on the splitting of the Coulomb electron-electron interaction (r_{12}^{-1}) into a short-range and a long-range part following the general scheme introduced by Savin [192]. The partition of the electron-electron repulsion is made in the coordinate space, commonly by the use of the error function and the introduction of a range parameter μ that controls the extent of the separation:

$$\frac{1}{r_{12}} = \frac{\mathrm{erfc}(\mu r_{12})}{r_{12}} + \frac{\mathrm{erf}(\mu r_{12})}{r_{12}}, \tag{4.1}$$

where the short- and long-range contributions are described respectively by the first and second terms of the RHS of Eq. (4.1). Such decomposition leads to energy functionals which involve a μ-dependent many-body wave function Ψ^μ, with an electron-electron interaction that is softer than r_{12}^{-1} for small r_{12}.

The non-relativistic electronic Hamiltonian for a molecular system can be written as:

$$\hat{H} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{sr,\mu} + \hat{V}_{ee}^{lr,\mu}, \tag{4.2}$$

in which $\hat{T}$, $\hat{V}_{ne}$ correspond to the operators of kinetic energy and nuclei-electron attraction. The corresponding parts of the electron-electron repulsion interaction $\hat{V}_{ee} = \hat{V}_{ee}^{sr,\mu} + \hat{V}_{ee}^{lr,\mu}$ (superindex μ explicitly indicates the dependence on the μ-parameter) are determined by the range separation parameter. By taking into account the effect of the rest of electrons, the long-range part of this last operator is given by:

$$\hat{V}_{ee}^{lr,\mu} = \sum_{i<j} \frac{\mathrm{erf}(\mu r_{ij})}{r_{ij}}, \tag{4.3}$$

where $r_{ij} = |r_i - r_j|$. This operator is applied to a multiconfigurational wave function, while the short-range interaction is treated by DFT. In the $\mu \to 0$ limit, the model converges

to the Kohn-Sham scheme, whereas for $\mu \to \infty$ the model recovers the WFT method with the full $1/r_{ij}$ interaction. Within this approach, the general expression for the energy is given by

$$\begin{aligned} E^{\text{WFT-}sr\text{DFT}} = \min_{\Psi^\mu} \langle \Psi^\mu | \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{lr,\mu} | \Psi^\mu \rangle \\ + E_H^{sr,\mu}[\rho] + E_x^{sr,\mu}[\rho] + E_c^{sr,\mu}[\rho], \end{aligned} \tag{4.4}$$

where $\rho \equiv \rho[\Psi^\mu]$, and $E_H^{sr,\mu}[\rho]$, $E_x^{sr,\mu}[\rho]$ and $E_c^{sr,\mu}[\rho]$ are the short-range Hartree, exchange, and correlation energy functionals, respectively.

In practice, application of Eq. (4.4) requires explicit expressions for the short-range energy functional contributions. The short-range Hartree energy can be written in terms of the one-particle density matrix (D_{pq}) and the short-range part of the two-electron interaction integral ($g_{pqrs}^{sr,\mu}$) as:

$$E_H^{sr,\mu}[\rho] = \frac{1}{2} \sum_{pq,rs} D_{pq} g_{pqrs}^{sr,\mu} D_{rs} = \frac{1}{2} \sum_{pq} D_{pq} j_{pq}^{sr,\mu} \tag{4.5}$$

$$\rho(\mathbf{r}) = \sum_{pq} D_{pq} \phi_p^*(\mathbf{r}) \phi_q(\mathbf{r}), \tag{4.6}$$

$$g_{pqrs}^{sr,\mu} = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_p^*(\mathbf{r}_1) \phi_q(\mathbf{r}_1) \frac{1 - \text{erf}(\mu r_{12})}{r_{12}} \phi_r^*(\mathbf{r}_2) \phi_s(\mathbf{r}_2). \tag{4.7}$$

The exchange and correlation functionals can be built as short-range versions of local spin-density (LSDA) or gradient generalized (GGA) approximations. For the correlation part, we use the short-range Vosko-Wilk-Nusair (*sr*VWN) [305], the Perdew-Wang (*sr*PW92) [306] and the Perdew-Burke-Ernzerhof (*sr*PBE) [300, 80]. For the exchange term we consider the short-range Slater-Dirac (*sr*LSDA) [307] and the PBE exchange (*sr*PBE) [300, 80] to be respectively combined with the short-range LSDA and GGA correlation functionals.

4.3.2 State average approximation

One important feature of the WFT-*sr*DFT scheme is that the expression for the WFT-*sr*DFT energy is state specific, the terms $E_H^{sr,\mu}[\rho]$ and $E_{xc}^{sr,\mu}[\rho]$ are non-linear in the one-electron density matrix. Thus the diagonalization of the Hamiltonian matrix is not allowed in the presence of several eigenstates. To avoid such non-linearity one can express electronic state densities in terms of a reference density matrix ($\mathbf{D}^{ref}$) [308], from which the exchange-correlation potential is evaluated. The expression for the Hartree and exchange-correlation terms in Eq. (4.4) (after an expansion around $\mathbf{D}^{ref}$), using matrix notation can be approximated as [267]:

$$\begin{aligned} E_{Hxc}^{sr,\mu} \approx E_{xc}^{sr,\mu}[\mathbf{D}^{ref}] - \frac{1}{2}\text{Tr}[\mathbf{j}^{sr,ref}\mathbf{D}^{ref}] \\ -\text{Tr}[\mathbf{v}_{xc}^{sr,ref}\mathbf{D}^{ref}] + \text{Tr}[(\mathbf{j}^{sr,ref} + \mathbf{v}_{xc}^{sr,ref})\mathbf{D}] \end{aligned} \tag{4.8}$$

where, $\mathbf{j}^{sr,ref}$ is the matrix form of the Hartree potential and $\mathbf{v}_{xc}^{sr,ref}$ is the exchange-correlation potential matrix, both obtained from the reference density. The density matrix in Eq. (4.8) can be defined as the average density over *N* electronic states.

$$\mathbf{D}^{ref} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{D}^{(i)}. \tag{4.9}$$

Taking into account this average density [Eq. (4.9)], we can express an effective Hamiltonian [Eq. (4.10)] from which the WFT-*sr*DFT wave functions and energies are defined as their eigenstates and eigenvalues [Eq. (4.11)].

$$\hat{H}^{\mu,\text{eff}} \equiv \hat{T} + \hat{V}_{ne} + \hat{V}_{ee}^{lr,\mu} + \hat{V}_{ee}^{sr,\mu,ref} \tag{4.10}$$

$$\hat{H}^{\mu,\text{eff}} \left|\Psi^{\mu}\right\rangle = E^{\mu,\text{eff}} \left|\Psi^{\mu}\right\rangle \tag{4.11}$$

Application of Eq. (4.8) has to be done with caution since the use of a state-averaged 1-PDM affects the variationality of the computed wave functions. It is also worth noting that the use of a reference density implies that a fraction of the short-range DFT contribution is constant and equal for all states (same reference density). This fraction corresponds to the last three terms in Eq. (4.12). The final expression for the WFT-*sr*DFT energy reads as

$$\begin{aligned} E^{\text{WFT-}sr\text{DFT}} &= E^{\mu,\text{eff}} + E_{xc}^{sr,\mu}[\mathbf{D}^{ref}] \\ -\frac{1}{2}\text{Tr}[\mathbf{j}^{sr,ref}\mathbf{D}^{ref}] &- \text{Tr}[\mathbf{v}_{xc}^{sr,ref}\mathbf{D}^{ref}]. \end{aligned} \tag{4.12}$$

4.3.3 From paired to unpaired electrons

In order to test the validity of WFT-*sr*DFT in the characterization of open-shell systems, we explore its performance in the description of the H_2 molecule potential energy surface. The ground state energy profile along a single bond dissociation, such in the H_2 stretching, can be used as a toy model of the conversion of a closed-shell structure (at the equilibrium bond distance) to an open-shell singlet with two unpaired electrons, *i.e.*, diradical (at dissociation), through an intermediate diradical character region. This simple system has been repeatedly employed to test electronic structure methods. Near the equilibrium, monoconfigurational approaches like Hartree-Fock and Kohn-Sham provide a good description of the ground state electronic structure [309], with two electrons in the σ-bonding orbital, but fail at long interatomic distances. As the molecule is stretched, the σ and σ^* orbitals become closer in energy and the ground state wave function, with partial electron occupancy of both σ and σ^*, is no longer well described by a single Slater determinant [309]. The correct wave function at the dissociation limit takes the form given by Heitler-London [310], in which one electron lie in each of the two hydrogen atoms.

Figure 4.1 shows the H_2 energy profile computed with full CI (FCI) and the 6-31G basis compared to RAS-SF and RAS-*sr*LDA, both using the ROHF $\sigma^1\sigma^{*1}$ high-spin triplet as the reference configuration and with the two electrons and two orbitals forming the RAS2 space.

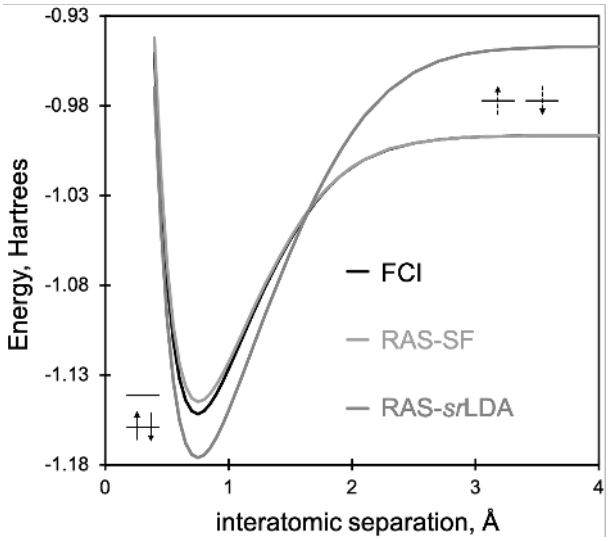


Figure 4.1: Ground state dissociation energy curves of H_2 computed at the FCI, RAS-SF and RAS-*sr*LDA levels with the 6-31G basis set.

In this case, the H_2 ground state potential energy dissociation profile computed at the RAS-SF level is in very good agreement with the FCI solution. The largest difference is obtained at the energy minimum, with the RAS-SF wave function lying 0.8 mHartrees above the FCI energy. This energy difference can be attributed to the lack of dynamic correlation in RAS-SF. Hence, one would expect RAS-*sr*DFT to improve the RAS-SF profile by slightly lowering the ground state energy in the equilibrium region. Indeed, the range-separated RAS-*sr*LDA ($\mu = 0.4$ a.u.) energy at the bonding distance is lower than RAS-SF (and FCI), but the overall shape of the PES has deteriorated considerably. The relative energy between closed- and open-shell forms, *i.e.*, binding energy, is largely overestimated by RAS-*sr*LDA. Nearly equal results are obtained with other short-range functionals and reference states (See resulting PES from Figure 4.3), suggesting that RAS-*sr*DFT as presented above might not be adequate for the study of open-shell systems, in particular for the relative energies of closed- and open-shell electronic structures, *e.g.*, different regions of molecular PESs or electronic states with different electronic nature.

In order to understand the origin of the large energy difference between the PES minimum and the dissociation limit in RAS-*sr*DFT we compute the RAS-*sr*LDA energy profile of the $M_S = 0$ and $M_S = \pm 1$ triplets (Figure 4.2a) and their exchange and correlation contributions (Figure 4.2b).

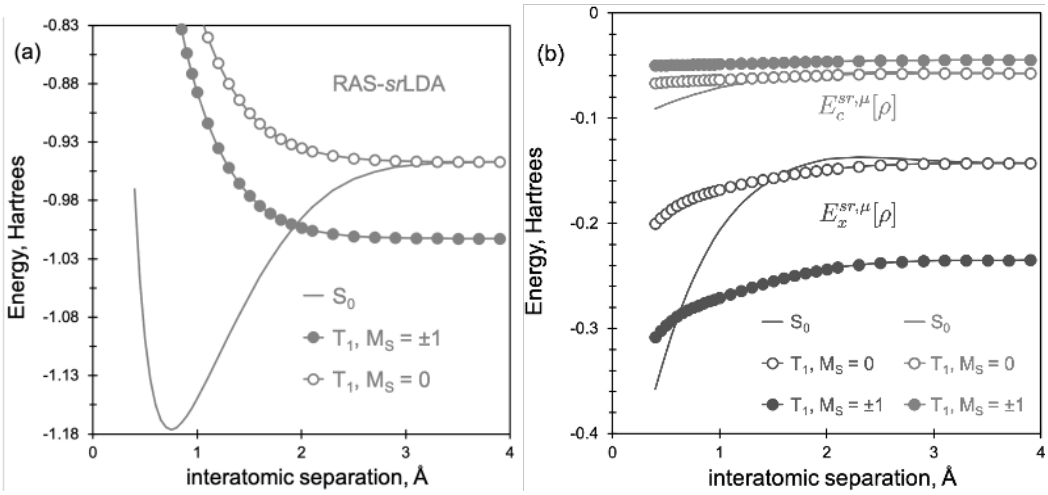


Figure 4.2: RAS-*sr*LDA dissociation energy curves of H_2 for the ground state singlet (S_0) and the T_1 with $M_S = 0$ and $M_S = \pm 1$ triplet microstates (a) and the *sr*-DFT exchange and correlation contributions (b) computed with the 6-31G basis set.

Importantly, RAS-*sr*LDA does not preserve the degeneracy between triplet components, *i.e.*, $M_S = 0$ and high/low-spin ($M_S = \pm 1$) microstates (Figure 4.2a). The $M_S = 0$ triplet is higher in energy than the degenerated $M_S = \pm 1$ components, and converges to the ground state singlet energy at long interatomic separation. Moreover, we notice that the dissociation limit for the $M_S = \pm 1$ would result in a much improved comparison with FCI profile and binding energy. On the other hand, representation of the exchange and correlation *sr*DFT energies shows that the magnitude of the exhange component is much larger than the short-range correlation energy contribution (Figure 4.2b), which is rather constant for the triplet state while it becomes more negative for the S_0 at short distances. The exchange energy exhibits a stronger dependence on interatomic separation, especially for the singlet state. Besides, the correlation energy of $M_S = 0$ and $M_S = \pm 1$ triplets is rather close to each other, whereas the difference between exchange energy is much larger (in the order of 0.1 Hartree) and dominates the gap between the triplet components in Figure 4.2a.

Therefore, we deduce that the accurate description of the H_2 dissociation with RAS-*sr*DFT, and the balanced characterization of closed and open-shell systems in general, might need to include short-range spin polarization effects, which have an important impact on the exchange energy. To that aim, in the following we consider different strategies directly related with spin polarization of the electron density, like in the conventional density functionals employed in (singlet determinant) unrestricted Kohn-Sham (UKS), and by exploring the use of exact exchange.

4.3.4 Spin polarized short-range DFT

Within the KS-DFT scheme, energy functionals incorporate spin polarization effects related with the presence of electron unpaired densities by breaking the spatial symmetry between spin-up (α) and spin-down (β), that is, unrestricted DFT (UDFT), through the spin polarization parameter (ζ) defined as $(\rho_\alpha - \rho_\beta)/\rho$ at each point in space. On the other hand, electron densities for $M_S = 0$ ($n_\alpha = n_\beta$) states derived from multiconfigurational wave functions, exhibit the same spatial distribution of α and β electrons ($\zeta = 0$). Hence, spin polarization

effects do not emerge in WFT-*sr*DFT. To include them, we employ the correlation functional scheme introduced by Pérez-Jiménez and collaborators [311], and based on the definition of two new densities:

$$\rho_{>} = \sum_{i,n_i \geq 1} (n_i - 1)|\phi_i|^2 \tag{4.13}$$

$$\rho_{<} = \rho - \rho_{>}, \tag{4.14}$$

where ϕ_i is the *i*th natural orbital with electron occupancy n_i. Then, the short-range exchange and correlation functionals are defined as $E_x^{sr,\mu}[\rho] \equiv E_x^{sr,\mu}[\rho_{>}, \rho_{<}]$ and $E_c^{sr,\mu}[\rho] \equiv E_c^{sr,\mu}[\rho_{>}, \rho_{<}]$, with:

$$\zeta = \frac{\rho_{>}(\mathbf{r}) - \rho_{<}(\mathbf{r})}{\rho(\mathbf{r})}. \tag{4.15}$$

We therefore label this methodology as WFT-*sr*-*p*DFT (or simply *sr*-*p*DFT). It is important to notice that we employ both $\rho_{>}$ and $\rho_{<}$ densities in the correlation and exchange functionals. Also, the state average scheme in Eqs. (4.8)-(4.12) should not be used with the definition of the $\rho_{<}$ and $\rho_{>}$ densities, since now the reference density ($\mathbf{D}^{ref}$) and the density of each state $\mathbf{D}$ might be rather different. Therefore, the spin polarization approach should be used in state-specific calculations.

4.3.5 Exact short-range exchange

In addition, here we propose two new approximations of $E_x^{sr,\mu}[\rho]$ based on the exact form of the (short-range) exchange interaction. First, we consider the Hartree-Fock exchange functional (*sr*EXX or *sr*-EXX):

$$E_x^{sr\text{EXX},\mu}[\rho] = -\frac{1}{2} \sum_{pq,rs} D_{pq} g_{psqr}^{sr,\mu} D_{rs} \tag{4.16}$$

Of course, Eq. (4.16) can be generalized to the case of short-range hybrid functionals in which $E_x^{sr\text{EXX},\mu}$ is mixed with *sr*DFT exchange:

$$E_x^{sr,\mu}[\rho] = a E_x^{sr\text{EXX},\mu}[\rho] + (1-a) E_x^{sr\text{DFT},\mu}[\rho], \tag{4.17}$$

where a is the mixing parameter.

Alternatively, we approximate the short-range exchange to the difference between the expectation value of the $\hat{V}_{ee}^{sr,\mu}$ and the short-range Hartree energy:

$$E_x^{sr\text{WFT},\mu}[\rho] \approx \min_{\Psi^\mu} \langle \Psi^\mu | \hat{V}_{ee}^{sr,\mu} | \Psi^\mu \rangle - E_H^{sr,\mu}[\rho], \tag{4.18}$$

which we name *sr*WFT exchange. Inserting Eq. (4.18) in the WFT-*sr*DFT energy [Eq. (4.4)] results in:

$$E^{\text{WFT-}sr\text{DFTc}} = \min_{\Psi^\mu} \langle \Psi^\mu | \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} | \Psi^\mu \rangle + E_c^{sr,\mu}[\rho], \tag{4.19}$$

Since Eq. (4.19) does not contain short-range DFT exchange, we name it WFT-*sr*DFTc, where the *c* stands for correlation. This approach assumes that $E_x^{sr\text{WFT},\mu}[\rho]$, concretely the first term on the RHS of Eq. (4.18), does not include short-range correlation. We might predict it to be a reasonable approximation for multiconfigurational wave functions with a limited number of configurations, *e.g.*, CASSCF with a small active space or small RAS2 in RASCI, while increasingly large number of correlated orbitals and electrons might induce double

counting of electron correlation. Such situation becomes evident for FCI wave functions, although the WFT term still depends on the μ parameter through the wave function.

We further approximate Eq. (4.19) by using the WFT density to evaluate the short-range DFT correlation energy for each state, that is, by avoiding the self-consistent iteration procedure required in WFT-*sr*DFTc. We define this "one-shot" energy correction as WFT-*sr*DFTc0:

$$E_i^{\text{WFT-}sr\text{DFTc0}} = E_i^{\text{WFT}} + E_c^{sr,\mu}[\rho_i]. \tag{4.20}$$

Such approximation can be justified by the fact that, although (*sr*DFT) electron correlation can have a sensible impact on electronic energies, in general, the converged WFT-*sr*DFTc densities are expected to mildly deviate from the WFT ones. Other similar simplifications and approximations have been described in the literature, such as in the use of the HF density as input for the evaluation of DFT exchange-correlation energy [312], or in the one-shot GW approximation (or G_0W_0) employed in the framework of Green's function-based methods [140, 313]. Importantly, the state specific one-shot DFT correlation energy in WFT-*sr*DFTc0 does not increase the cost of WFT. Moreover, the state average approximation (section 4.3.2) is no longer needed in the computation of excitation energies.

4.4 Computational Details

In this Chapter we use the RASCI within the *hole* and *particle* approximation wave function [304] in combination with short-range DFT functionals. Specific details of the reference configuration, RAS2 orbital space and excitation operators employed are explicitly indicated in each example. In all cases, RAS1 (RAS3) orbital set contain all fully occupied (virtual) orbitals below (above) the RAS2 space. Unless indicated, in all the examples of the present study we have used the short-range Slater-Dirac (*sr*LSDA) [307] functional for the exchange term and the short-range version of the Vosko-Wilk-Nusair functional (*sr*VWN) [314] for the correlation part. Previous calculations have shown weak dependence with the short-range functional [267]. As pointed previously by Gori-Giorgi [76], if the range-separation parameter (μ) is not too large, few determinants describe Ψ^μ quite accurately. Previous studies have suggested the optimal value for the μ-parameter to be within the $0.4 \leq \mu \leq 0.5$ a.u. range [202, 315, 267]. Therefore, here we systematically use $\mu = 0.4$ a.u. All calculations presented in section 4.5 were obtained with a development version of the Q-Chem package [316].

4.5 Numerical results

In the following we present several examples of the performance of the introduced methods in comparison to highly accurate calculations and experimental measurements. First of all, we explore the capabilities of the WFT-*sr*DFT methods in the single bond breaking of the H_2 molecule and energy gaps in triplet atoms (C, O, Si, and S), in triatomic diradicals (CH_2, NH_2^+, SiH_2 and PH_2^+, in benzyne isomers, and in trimethylenemethane. Finally, we apply the new methods to the calculation of low-lying electronic states of Ga_2 dimer. Complementary information for H_2 (number of unpaired electrons) and triatomic diradicals (additional adiabatic ST gaps) can be found in Appendix A.

4.5.1 Hydrogen dissociation

First, we revisit the H_2 PES, Figure 4.3 presents calculations exploring the dependence with the specific form of the short-range DFT functional and different values of the range separation parameter (μ). As observed, different forms of the exchange-correlation *sr*-density functional (*sr*-DFT) give similar energy profiles within an increase in the energy at dissociation range and a decrease at bonding range. When fixing the exchange-correlation and changing μ further changes illustrate the effect of going from the WFT dominant solution ($\mu \to \infty$) to the DFT dominant solution ($\mu \to 0$), changes in the PES mainly occur at dissociation range. These are general features displayed by the RAS-*sr*DFT approach.

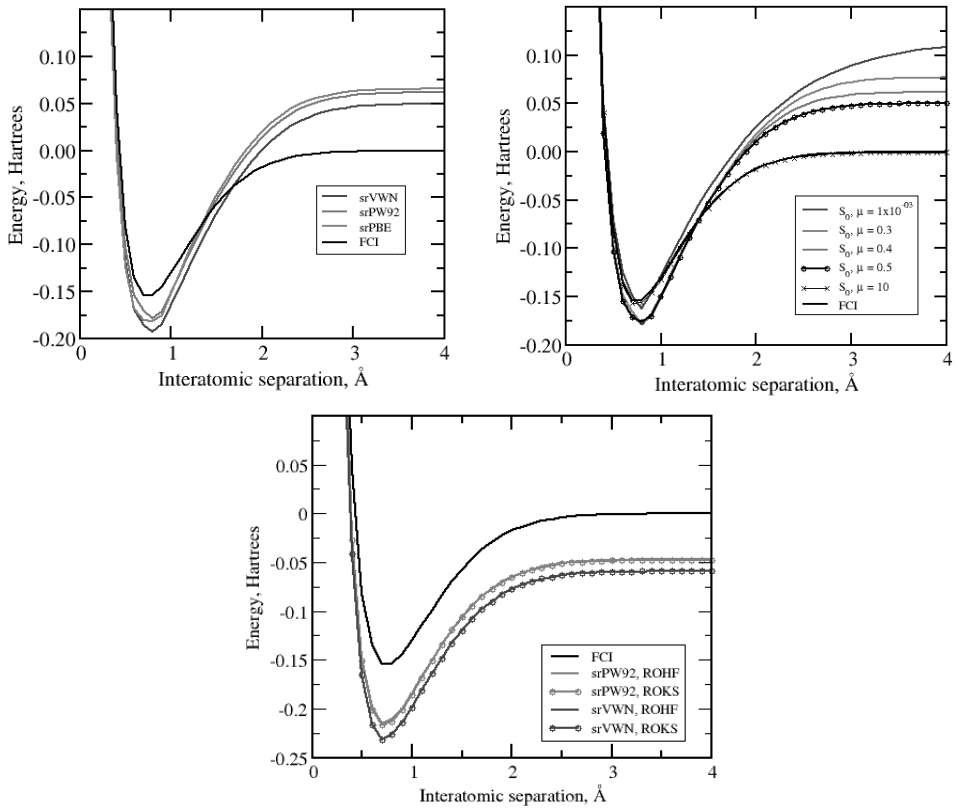


Figure 4.3: RAS-*sr*-DFT dissociation energy curves of H_2 for the ground state singlet (S_0) computed with different short-range DFT functionals (upper left plot), with RAS-*sr*-LDA and different values for the range separation parameter μ (upper right plot). And computed with different *sr*DFT functional and different high-spin triplet reference configurations (ROHF and ROKS) (lower central plot). The Slater exchange functional has been employed for the ROKS reference, while the correlation functional corresponds to the full range counterpart to the one used in the *sr*DFTc calculation, *i.e.*, VWN or PW92. All calculations have been done with the 6-31G basis set.

Calculations with the newly introduced short-range functionals *sr-p*LDA, *sr*-EXX, *sr*-LDAc and *sr*-LDAc0 are shown in Figure 4.4. The spin polarization approximation is able to stabilize the S_0 (and the $M_S = 0$ triplet) energy for large interatomic separations, which actually converges to the *sr*-LDA $M_S = \pm 1$ triplet energy (Figure 4.4, right plot). As a result, the computed binding energy notably reduces with respect to *sr*-LDA. But it does not produce the right energy profile at intermediate distances, with an unphysical energy maximum at ~ 2.5 Å. Replacing the short-range Slater exchange by the short-range part of the Hartree-

Fock exchange does not improve the *sr*-LDA result, with an even larger overestimation of the binding energy. On the other hand, the shape of the short-range correlation only (*sr*-LDAc) approach is considerably improved with respect to the other *sr*DFT schemes, with a PES rather parallel to FCI. Ground state H_2 energy along the molecular dissociation obtained with the one-shot *sr*-LDAc (*sr*-LDAc0) are almost exactly equal to *sr*-LDAc, and are not represented in Figure 4.4 for the sake of clarity.

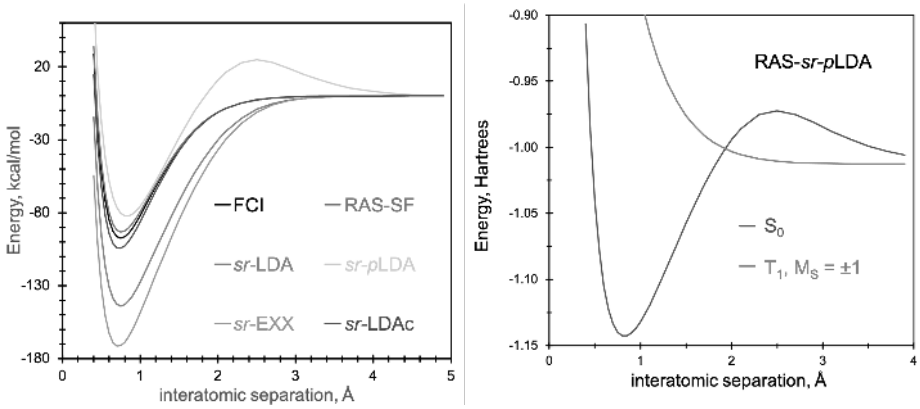


Figure 4.4: Ground state dissociation energy curves (in kcal/mol) of H_2 with respect to the dissociation limit computed at the FCI, RAS-SF, RAS-*sr*LDA (*sr*-LDA), RAS-*sr*LDA with with spin polarized densities (*sr*-*p*LDA), with short-range HF exchange (*sr*-EXX), and with short-range correlation (*sr*-LDAc) levels (left plot). And RAS-*sr*-*p*LDA dissociation energy curves of H_2 for the ground state singlet (S_0) and the T_1 with $M_S = \pm 1$ triplet microstates (right plot). Calculations done with the 6-31G basis set.

4.5.2 Triplet state atoms: C, O, Si and S

Next, we explore the performance of RAS-*sr*DFT methods in the computation of triplet-singlet energy gaps in carbon, oxygen, silicon and sulfur atoms. All four atoms hold a triply degenerated spin-triplet ground state (^3P) and a five-fold first excited singlet (^1D). For a balanced description of the ^3P and ^1D multiplets, we employ different excitation schemes. For carbon and silicon we use a SF operator with a high-spin quintet reference and an IP operator with a high-spin quartet anion configuration. For oxygen and sulfur we combine an IP operator with the dianion singlet configuration and an AE operator with the high-spin quartet cation reference determinant. The RAS2 space is built with four electrons in four orbitals (ns^2np^2) for carbon ($n = 2$) and silicon ($n = 3$) atoms, and with six electrons in four orbitals (ns^2np^4) for oxygen ($n = 2$) and sulfur ($n = 3$). Tables 4.1 and 4.2 show the energy gaps computed for the $M_S = 0$ (ΔE_0) and $M_S = \pm 1$ (ΔE_1).

The different versions of the pure RASCI method (no short-range DFT) preserve spatial and spin orientation degeneracies, that is, the ^3P state energy is invariant with respect to M_L and M_S (nine degenerated microstates), and the ^1D state is degenerated for $-2 \leq M_L \leq 2$ (five degenerated singlets). As a result, the computed triplet-singlet gaps are invariant with respect to the triplet state projection, *i.e.* $\Delta E_0 = \Delta E_1$ in Tables 4.1 and 4.2. On the other hand, RAS-*sr*DFT methods are invariant with respect to spatial projections (M_L), but produce different energies for $M_S = 0$ and $M_S = \pm 1$ triplets ($\Delta E_0 \neq \Delta E_1$), as shown in Figure 4.2a for the H_2 molecule.

The four studied atoms are purely open-shell systems, and RASCI with the proper reference configuration, excitation operator (RAS-SF, RAS-IP and RAS-AE) and orbital space combination is expected to quantitatively describe the relative energies between the low-lying states (^3P and ^1D). Hence, ideally, we would like RAS-*sr*DFT to maintain the same accuracy in the computation of triplet-singlet energies, or improve it if possible. RASCI results with small RAS2 spaces are in good agreement with experimental gaps, and present small variations between the different considered reference configurations and excitation operators.

The RAS-*sr*LDA approach produces too small energy differences between the $M_S = 0$ triplet and ^1D, with triplet and singlet states nearly degenerated. The results improve considerably for the $M_S = \pm 1$ triplets, with rather accurate gaps for carbon (RAS-SF) and silicon, but still far from the RASCI accuracy for oxygen and sulfur. The polarized version of *sr*-LDA in general improves the results for the $M_S = 0$ case, but the gaps are still too small, especially for oxygen and sulfur. The results with *sr*-EXX tend to largely over stabilize the spin-singlet ^1D state with respect to ^3P, with too small (or even negative) triplet-singlet gaps. On the other hand, the short-range correlation approach (RAS-*sr*LDAc) produces rather accurate gaps, very close to RASCI values, and with slightly improved energies for the $M_S = \pm 1$ triplets. In general, polarized density RAS-*sr*LDAc calculations yield triplet-singlet gaps similar to those obtained without polarizing the electron density, or slightly closer to experiment in some cases.

From these results and the H_2 dissociation profiles, we might conclude that the straight use of RAS-*sr*DFT and WFT-*sr*DFT methodologies in general might not be a reliable approach for the study of open-shell systems. Amongst the different alternatives considered to overcome such limitation, the combination of multiconfigurational wave function (RASCI) exhange with short-range DFT correlation (WFT-*sr*DFTc and WFT-*sr*DFTc0) yields good relative energies between low-lying states, while the use short-range Hartree-Fock exchange fails to quantify singlet-triplet gaps and the relative energy between open-shell and closed-shell configurations. On the other hand, despite that the polarized density approach defined in Eqs. (4.13)-(4.15) yields good energy gaps, it prevents the use of the state average approximation, that is, state specific calculations are mandatory. Hence, it might be a good alternative when convergence to different individual electronic states can be easily achieved, much like ΔSCF and ΔKS strategies employed in the calculation of singlet-triplet gaps or ionization potentials [15, 317], but it might pose serious difficulties in the computation of transition energies to several excited states. Therefore, in the following sections we focus our attention on the performance of WFT-*sr*DFTc in comparison to pristine WFT and to WFT-*sr*DFT, that is, short-range DFT exchange and correlation.

Table 4.1: Energy gaps (in eV) between the ground state triplet (^3P) and the lowest singlet state (^1D) in carbon and silicon atoms, computed with RAS-SF (high-spin quintet as the reference configuration) and RAS-IP (high-spin quartet anion as the reference configuration) with the pristine wave function (RASCI) and different implementations of RAS-*sr*DFT and the cc-pVTZ basis set. ΔE_0 and ΔE_1 values correspond to energy gaps with respect to the $M_S = 0$ and $M_S = \pm 1$ ^3P triplet, respectively. Experimental gaps indicated in parenthesis.

	ΔE_0		ΔE_1		ΔE_0		ΔE_1	
method	SF[a]	IP[b]	SF	IP	SF	IP	SF	IP
	Carbon (exp. 1.26)				Silicon (exp. 0.75)			
RASCI	1.50	1.49	1.50	1.49	0.98	0.97	0.98	0.97
sr-LDA	0.08	0.08	1.27	-1.75	0.18	0.18	0.81	0.81
sr-*p*LDA	1.23	1.05	1.11	0.90	0.55	0.16	0.38	0.00
sr-EXX	0.09	0.10	-0.38	-0.37	0.20	0.21	-0.43	-0.42
sr-LDAc	1.55	1.51	1.18	1.14	1.01	0.99	0.90	0.88
sr-LDAc0	1.56	1.34	1.28	1.13	1.01	0.99	0.88	0.86
sr-*p*LDAc	1.30	1.36	1.46	1.37	0.54	0.91	1.16	0.92

[a]SF label stands for RAS-SF method. [b]IP label stands for RAS-IP method.

Table 4.2: Energy gaps (in eV) between the ground state triplet (^3P) and the lowest singlet state (^1D) in oxygen and sulfur atoms, computed with RAS-IP (singlet dianion as the reference configuration) and RAS-AE (high-spin quartet cation as the reference configuration) with the pristine wave function (RASCI) and different implementations of RAS-*sr*DFT and the cc-pVTZ basis set. ΔE_0 and ΔE_1 values correspond to energy gaps with respect to the $M_S = 0$ and $M_S = \pm 1$ ^3P triplet, respectively. Experimental gaps indicated in parenthesis.

	ΔE_0		ΔE_1		ΔE_0		ΔE_1	
method	IP[a]	AE[b]	IP	AE	IP	AE	IP	AE
	Oxygen (exp. 1.97)				IP[a]	AE[b]	IP[a]	AE[b]
RASCI	2.05	2.16	2.05	2.16	1.29	1.34	1.29	1.34
sr-LDA	0.04	0.03	1.44	1.50	0.13	0.12	0.85	0.86
sr-*p*LDA	1.18	1.64	0.91	1.44	0.64	0.82	0.39	0.57
sr-EXX	-0.19	-0.21	-0.30	-0.31	0.12	0.12	-0.20	-0.20
sr-LDAc	2.06	2.19	2.00	2.16	1.30	1.36	0.99	1.07
sr-LDAc0	2.07	2.20	1.73	1.84	1.30	1.36	1.11	1.16
sr-*p*LDAc	1.99	1.98	1.78	1.97	0.99	1.24	1.26	1.23

[a]IP label stands for RAS-IP method. [b]EA label stands for RAS-EA method.

4.5.3 Triatomic diradicals: CH_2, NH_2^+, SiH_2 and PH_2^+

Methylene (CH_2), nitrenium cation (NH_2^+), silylene (SiH_2) and the dihydridophosphorus(1+) cation (PH_2^+) are four isovalent molecules classified by Salem [275] as heterosymmetric diradicals. For such species, the valence orbitals ϕ_1 and ϕ_2 are derived from a p-orbital and a $s^{1-x}p^x$ hybridized orbital of the heavy atom. The low-lying electronic states of these four triatomic diradicals are obtained as the different distribution of two valence electrons in the two orbitals to form a triplet state ($^3B_1$) and three singlets: two closed-shell states with $^1A_1$ symmetry and the open-shell $^1B_1$ state. The gap between the valence orbitals gradually changes in the series, giving as a consequence various degrees of diradical character. The ground state of CH_2 and NH_2^+ is the triplet $\widetilde{X}\,^3B_1$ state, with the three excited singlets energetically ordered (from lower to higher energy) as: $\widetilde{a}^1A_1$, $\widetilde{b}^1B_1$ and $\widetilde{c}^1A_1$. The latter

formally being a double excitation from the $\widetilde{a}^1A_1$ state. On the other hand, SiH_2 and PH_2^+ molecules have a closed-shell singlet ground state ($\widetilde{X}^1A_1$) followed by the open-shell singlet $\widetilde{A}^1B_1$, the triplet $\widetilde{a}^3B_1$ and the singlet $\widetilde{B}^1A_1$ as their lowest excited states.

Table 4.3: Adiabatic energy gaps (in eV) with respect to the $\widetilde{X}^3B_1$ ($M_s = 1$) ground state for CH_2 and NH_2^+ molecules computed with RAS(n, m)-SF and RAS(n, m)-SF-*sr*LDAc, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively.

method	$\widetilde{a}^1A_1$	$\widetilde{b}^1B_1$	$\widetilde{c}^1A_1$
CH_2			
RAS(2,2)-SF	0.595	1.656	3.257
RAS(6,6)-SF	0.648	1.661	3.246
RAS(2,2)-*sr*LDAc	0.344	1.359	3.012
RAS(6,6)-*sr*LDAc	0.371	1.365	2.989
RAS(6,6)-*sr*LDAc0	0.365	1.418	3.050
FCI[a]	0.483	1.542	2.674
Expt.[b]	0.390	1.425	
Expt. $-\Delta$ ZPE[c]	0.407	1.430	
NH_2^+			
RAS(2,2)-SF	1.424	1.992	3.957
RAS(6,6)-SF	1.543	2.114	3.912
RAS(2,2)-*sr*LDAc	1.349	1.865	3.880
RAS(6,6)-*sr*LDAc	1.463	2.005	3.834
RAS(6,6)-*sr*LDAc0	1.182	1.777	3.610
CASSCF SOCI[a]	1.281	1.935	3.380
Expt.[d]	1.306		
Expt. $-\Delta$ ZPE[e]	1.286		

[a] TZ2P basis from Ref. [318]. [b] From Ref. [230]. [c] ZPE changes at the CISD level from Ref. [319]. [d] From Ref. [320]. [e] ZPE changes at the CISD level from Ref. [321].

Computed adiabatic energy differences between the three singlets and the triplet state are shown in Table 1 (CH_2 and NH_2^+) and Table 4.4 (SiH_2 and PH_2^+). The performance of *sr*DFTc with a RASCI multiconfigurational wave functions obtained by an spin-flip (α-to-β) excitation operator acting on the ROHF high-spin triplet reference configuration is evaluated for two different RAS2 spaces: (*i*) two electrons in the two frontier orbitals, and (*ii*) six electrons in six orbitals, and compared to the pure wave function (RAS-SF) method. Molecular geometries for each state were taken from previous computational works, with CH_2 structures optimized at the FCI/TZ2P level [322], and NH_2^+, SiH_2 and PH_2^+ equilibrium geometries obtained with CISD/TZ2P(f,d) [321, 323, 324]. All RAS-SF and RAS-SF-*sr*DFTc calculations have been done with the cc-pVTZ basis set.

Pristine WFT excitation energies at the RAS-SF level are systematically higher than experimental values available and reference energies, *i.e.*, FCI and CASSCF SOCI for CH_2 and NH_2^+, respectively. The RAS-SF errors for the $\widetilde{a}^1A_1$ and $\widetilde{b}^1B_1$ states are within the 0.2-0.3 eV range, and considerably large for the transition to $\widetilde{c}^1A_1$ (> 0.5 eV). Short-range DFT correlation reduces the adiabatic excitation energies to the three singlets by a few tenths of an eV, which results in energy gaps closer to the reference values, that is, with absolute errors of the order of 0.1-0.3 eV for all cases. In general, for these molecules there is only a mild dependence of adiabatic energies with the size of the RAS2 correlated space for

both WFT and WFT-*sr*DFTc calculations. The *sr*DFTc0 energies are quite close to the ones obtained self-consistently with *sr*DFTc. It is worth noting that, in several cases one-shot computed gaps are even closer to the reference values that the ones obtained iteratively.

Table 4.4: Adiabatic ST gaps (in eV) with respect to the $\widetilde{X}^1A_1$ ground state for SiH_2 and PH_2^+ molecules computed with RAS(n,m)-SF and RAS(n,m)-SF-*sr*LDAc, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively.

method	$\widetilde{a}^3B_1$	$\widetilde{A}^1B_1$	$\widetilde{B}^1A_1$
SiH_2			
RAS(2,2)-SF	0.748	2.089	3.940
RAS(6,6)-SF	0.696	1.786	3.513
RAS(2,2)-*sr*LDAc	1.593	2.174	4.075
RAS(6,6)-*sr*LDAc	1.210	1.878	3.662
RAS(6,6)-*sr*LDAc0	0.929	1.919	3.699
CASSCF SOCI[a]	0.871	1.992	3.486
Expt.	0.91[b]	1.928[c]	
Expt. $-\Delta$ ZPE[d]	0.90	1.92	
PH_2^+			
RAS(2,2)-SF	0.616	2.121	4.287
RAS(6,6)-SF	0.531	1.778	3.837
RAS(2,2)-*sr*LDAc	1.038	2.185	4.415
RAS(6,6)-*sr*LDAc	0.978	1.871	3.980
RAS(6,6)-*sr*LDAc0	0.817	1.904	4.007
CASSCF SOCI[a]	0.760	2.009	3.686
Expt.[e]	0.75	1.92	
Expt. $-\Delta$ ZPE[f]	0.76	1.94	

[a] TZ2P basis from Ref. [318]. [b] From Ref. [325]. [c] From Ref. [326]. [d] ZPE changes at the CISD level from Ref. [323]. [e] From Ref. [327]. [f] ZPE changes at the CISD level from Ref. [324].

RAS-SF calculations for the SiH_2 and PH_2^+ diradicals tend to over stabilize the triplet $\widetilde{a}^3B_1$ state with respect to the three singlet states, like in the CH_2 and NH_2^+ molecules. Hence, computed adiabatic energy differences between the ground state singlet ($\widetilde{X}^1A_1$) and the $\widetilde{a}^3B_1$ state are lower than experimental and reference values. On the other hand, energy gaps to the two lowest excited singlets are in better agreement. It is worth noticing that in some cases the energies obtained with the smallest fully correlated space (two electrons in two orbitals) are more accurate than with the larger electron orbital space (six electrons in six orbitals), which might result from error compensation. The use of *sr*DFTc in combination with the RAS-SF wave function increases the computed gap to the lowest triplet, since short-range correlation is stronger in the singlet ground state. The adiabatic energies of the $\widetilde{A}^1B_1$ and $\widetilde{B}^1A_1$ states also increase with respect to the wave function values, but to a lesser degree. Once again, one-shot short-range correlation energies are, in general, slightly improved with respect to RAS-*sr*DFTc results.

Finally, it is worth noting that RAS-SF with short-range DFT exchange and correlation yields to large errors in the computation of adiabatic energies in all cases (CH_2, NH_2^+, SiH_2 and PH_2^+), with a systematic large over stabilization of the spin-singlet states with respect to the triplet (See Figure 4.5).

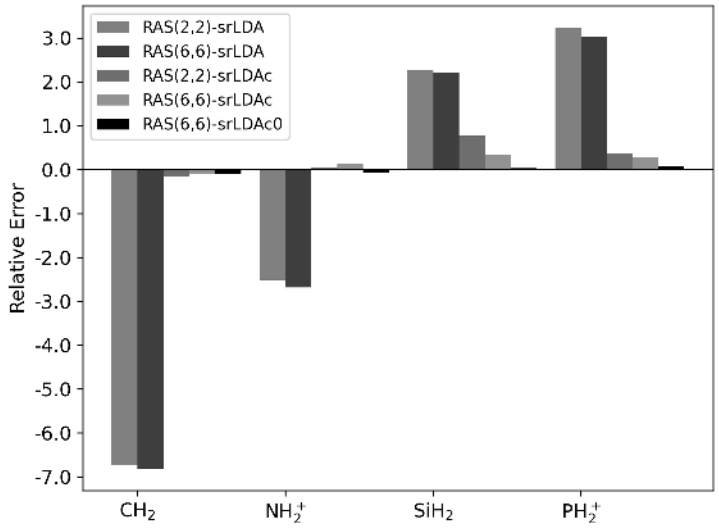


Figure 4.5: Relative error in triatomic diradical series. Errors correspond to the ST gaps with respect to the Expt. $-\Delta$ ZPE value reported from references as indicated in previous tables.

4.5.4 Benzyne isomers

The abstraction of two hydrogen atoms from benzene gives rise to prototypical aromatic diradical systems called didehydrobenzenes or benzynes (C_6H_4). Benzynes feature two unpaired electrons and exhibit a diradical character strength that increases with the distance between the carbons formally hosting the two unpaired electrons (*ortho* < *meta* < *para*). This is illustrated in Figure 4.6 where FOD and Head-Gordon indexes characterize the effective unparied electrons for for each case.

All three benzyne isomers present a singlet diradical ground state with an energy gap to the lowest triplet excited state, $^3B_2$ in *o*- and m-benzyne, and $^3B_{1u}$ in p-benzyne, that decreases from *ortho* to *meta* to *para* forms, that is, inversely proportional to the ground state diradical character [328]. The strong diradical character of benzynes, especially p-benzyne, leads to orbital instabilities in mean-field approximations, *e.g.*, (restricted) HF. As a consequence, the description of the electronic structure of benzynes requires the use of strongly correlated methods [318].

Here, we compute the adiabatic singlet-triplet energy gaps for the three benzyne isomers employing the RAS-SF method with and without short-range DFT correlation, defining a RAS2 space containing six electrons in the six frontier π-orbitals, and the cc-pVDZ basis (Table 4.5). In this case, we also explore the dependence of the computed singlet-triplet energies with the reference orbitals. To that aim, we consider the lowest triplet state obtained with the ROHF and ROKS (Slater exchange and PW92 correlation) schemes as the reference configuration. Equilibrium structures were optimized by Orms and co-workers at the SF-TDDFT (ground state singlet) and KS (lowest excited triplet) levels of theory, with the use of the B5050LYP functional and the cc-pVTZ basis set [255].

First of all, we notice that the RAS-SF wave function ansatz respects the energy degeneracy between the three components of the triplet state, resulting in identical ΔE_0 and ΔE_1 computed energies for both HF and KS reference configurations. On the other hand, the invariance of the (non-relativistic) electronic Hamiltonian with respect to the spin axis

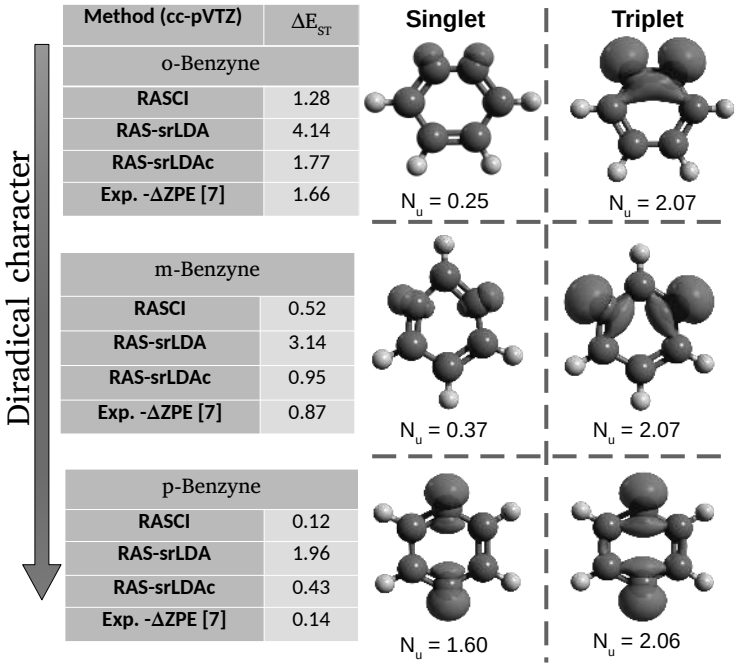


Figure 4.6: Accurate singlet-triplet energy gaps computed for benzyne isomers (in eV). Distribution of the unpaired electrons in the real space is correctly described by the fractional occupation density FOD.

orientation is not preserved by calculations including short-range DFT energy contributions, as becomes evident by the different values obtained regarding the employed triplet component. The magnitude of DFT correlation is larger for the $M_S = 0$ triplet than in $\alpha\alpha$ and $\beta\beta$ components, as shown in the H_2 dissociation (Figure 4.2b). Spin polarization effects are not taken into consideration in the correlation energy functional of the triplet microstate with the same number of spin-*up* and spin-*down* electrons, resulting in $|E_c^{sr,\mu}[\rho_0]| > |E_c^{sr,\mu}[\rho_{\pm 1}]|$.

Singlet-triplet energies obtained with ROHF and ROKS triplet orbitals do not present large differences, especially in the RAS-SF and RAS-SF-*sr*LDAc approaches (< 0.1 eV). In general, computed gaps are larger with the use of KS orbitals in o- and m-benzyne, while they are slightly lower in p-benzyne. In any case, these results indicate a weak dependence of the relative energies with the use of HF or KS reference configurations. RAS-SF results are in quite good agreement with the experimental values, with a systematic underestimation of the energy gaps, that is, in this case the spin-flip approach tends to over stabilize the triplet state with respect to the ground state singlet. RAS-SF with *sr*DFT exchange and correlation results in much to large singlet-triplet energies, in particular when the $M_S = \pm 1$ triplet is employed. These results confirm that WFT-*sr*DFT is not able to provide a balanced description between the open-shell triplet and the singlet diradicaloid. The use of the $M_S = 0$ triplet improves the results, but its performance is still far from the WFT (RAS-SF) results. Combining long- and short-range WFT exchange with *sr*DFTc increases the WFT (RAS-SF) adiabatic energy differences, as the (negative) short-range correlation energy is larger for the singlet state. The use of the high-spin triplet seems to be a better choice for the computation of singlet-triplet gaps in these diradicals, since it includes spin polarization effects in the triplet state energy, although the computed gaps for p-benzyne at the employed level of are closer to the experimental value when we take the $M_S = 0$ energy. In general, energies obtained for the three molecules with the *sr*DFTc0 approach are very close to the

*sr*DFTc counterparts.

Table 4.5: Singlet-triplet adiabatic energies (in eV) of benzyne isomers computed with RAS-SF, RAS-*sr*DFT and RAS-*sr*DFTc with two different RAS2 spaces, and with the ROHF (HF superindex in parenthesis) and ROKS (KS superindex in parenthesis) triplet reference configurations. Spin projection of the triplet state indicated with a subindex 0 ($M_S = 0$) or 1 ($M_S = \pm 1$).

method	$\Delta E_0^{(\mathrm{HF})}$	$\Delta E_0^{(\mathrm{KS})}$	$\Delta E_1^{(\mathrm{HF})}$	$\Delta E_1^{(\mathrm{KS})}$
		o-benzyne		
RAS(6,6)-SF	1.28	1.33	1.28	1.33
RAS(6,6)-*sr*LDA	2.30	2.30	4.14	4.08
RAS(6,6)-*sr*LDAc	1.42	1.48	1.77	1.82
RAS(6,6)-*sr*LDAc0	1.42	1.48	1.67	1.71
Expt.[a]		1.63		
Expt. $-\Delta$ ZPE[b]		1.66		
		m-benzyne		
RAS(6,6)-SF	0.52	0.60	0.52	0.60
RAS(6,6)-*sr*LDA	1.51	1.58	3.14	3.09
RAS(6,6)-*sr*LDAc	0.60	0.68	0.95	1.02
RAS(6,6)-*sr*LDAc0	0.61	0.68	0.85	0.91
Expt.[a]		0.91		
Expt. $-\Delta$ ZPE[b]		0.87		
		p-benzyne		
RAS(6,6)-SF	0.12	0.10	0.12	0.10
RAS(6,6)-*sr*LDA	0.49	0.39	1.96	1.87
RAS(6,6)-*sr*LDAc	0.12	0.11	0.43	0.43
RAS(6,6)-*sr*LDAc0	0.12	0.11	0.33	0.32
Expt.[a]		0.17		
Expt. $-\Delta$ ZPE[b]		0.14		

[a]From Ref. [329]. [b]ZPE computed at the SF-TDDFT/6-311G(d) level from Ref. [318].

4.5.5 Trimethylenemethane diradical

The trimethylenemethane (TMM) molecule is a fully conjugated non-kekulé diradical, that is, with all possible Kekulé structures containing at least two atoms that are not π-bonded [330]. In the ground state, TMM exhibits a D_{3h} flat structure with two electrons in the two degenerated e'' frontier π-orbitals (Figure 4.7). The electronic ground state corresponds to the spin-triplet configuration with single electron occupancy of the two e'' orbitals (ϕ_1 and ϕ_2 in Eqs. (4.21)-(4.26)), *i.e.*, the $^3A_2'$ state, in agreement with the Hund's rule and high level calculations [331, 332, 333, 334, 335, 336, 337, 318, 338].

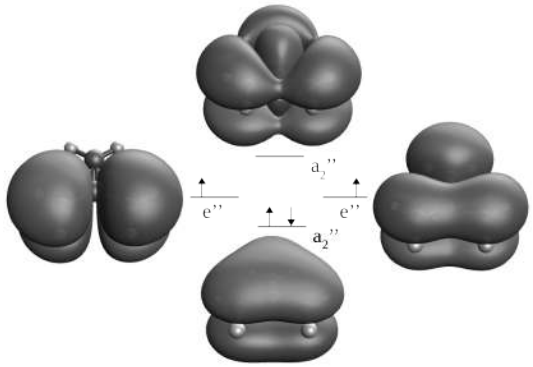


Figure 4.7: The π-system of TMM at the flat D_{3h} symmetry of the $^3A_2'$ ground state geometry computed at the ROHF/cc-pVTZ level.

The lowest lying electronic excited states at the $^3A_1$ PES minimum correspond to the symmetric spatial representation components of the direct product between the irreducible representation of the frontier orbitals ($E'' \otimes E''$), thus, they must be combined with anti-symmetric spin functions, *i.e.*, the $^1E'$ and $^1A_1'$ spin-singlet states. The doubly degenerated $^1E'$ state contains a closed-shell [Eq. (4.24)] and an open-shell [Eq. (4.25)] singlet, while the $^1A_1'$ is a closed-shell-like state mainly obtained as the in-phase combination of the doubly occupancy of the two e'' orbitals [Eq. (4.26)]. The two-fold $^3E'$ state, obtained as a single electron promotion from the occupied a_2'' orbital to the e'' orbital space, is the energetically lowest state beyond the four states generated by two electrons in the e'' orbital space.

$$\Psi_0(^3A_2') = \frac{1}{2}(\phi_1\phi_2 - \phi_2\phi_1)\cdot(\alpha\beta + \beta\alpha), \tag{4.21}$$

$$\Psi_1(^3A_2') = \frac{1}{\sqrt{2}}(\phi_1\phi_2 - \phi_2\phi_1)\cdot(\alpha\alpha), \tag{4.22}$$

$$\Psi_{-1}(^3A_2') = \frac{1}{\sqrt{2}}(\phi_1\phi_2 - \phi_2\phi_1)\cdot(\beta\beta), \tag{4.23}$$

$$\Psi_{CS}(E'') = \frac{1}{2}[(\phi_1)^2 - (\phi_2)^2]\cdot(\alpha\beta - \beta\alpha), \tag{4.24}$$

$$\Psi_{OS}(E'') = \frac{1}{2}(\phi_1\phi_2 + \phi_2\phi_1)\cdot(\alpha\beta - \beta\alpha), \tag{4.25}$$

$$\Psi(^1A_1') = \frac{1}{2}[(\phi_1)^2 + (\phi_2)^2]\cdot(\alpha\beta - \beta\alpha). \tag{4.26}$$

As for benzynes, the FOD representation enables us to characterize the unpaired electrons in TMM (see Fig. 4.8). As expected, the closed-shell $^1E''$ state displays zero effectively unpaired electrons ($N_u = 0$), and consequently, the FOD vanishes. For the remaining states, the FOD is non-vanishing and appears to be mostly localized over the C-C bonds of the molecule. Table 4.6 shows the computed vertical excitation energies from the $^3A_2'$ ground state of TMM in comparison with the previous results at the EOM-SF-CCSD. level [332] For this set of calculations we employ the reported ground state equilibrium geometry optimized at the SF-TDDFT/6-31G(d) level [318].

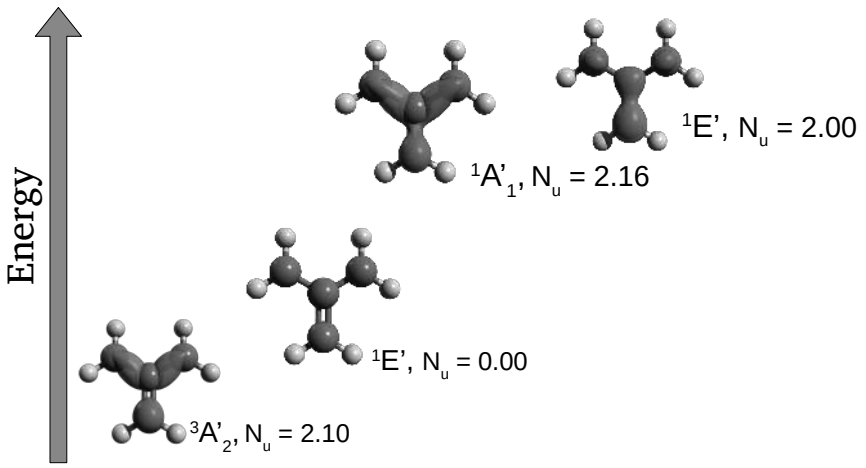


Figure 4.8: Low lying states of TMM biradical and respective FODs and effectively unpaired electrons accounted by the Head-Gordon's N_u index.

Table 4.6: Vertical excitation energies (in eV) to the low-lying excited states with respect to the ground $^3A'_2$ state of TMM computed at the RAS-SF and RAS-SF-*sr*LDAc (RAS-*sr*LDAc) with the cc-pVTZ basis set. The (n, m) values indicate the number of electrons (n) and orbitals (m) in the RAS2 space.

method	$^1E'$	$^1A'_1$	$^3E'$
RAS(4,4)-SF	1.13	5.09	4.61
RAS(4,4)-*sr*LDAc	0.99	5.18	4.60
RAS(4,4)-*sr*LDAc0	1.00	5.22	4.62
RAS(8,6)-*sr*LDAc	0.94	4.44	4.72
RAS(8,7)-*sr*LDAc	0.94	4.38	4.72
RAS(8,8)-*sr*LDAc	0.93	4.29	4.73
RAS(8,9)-*sr*LDAc	0.92	4.23	4.74
RAS(8,9)-*sr*LDAc0	0.94	4.27	4.74
EOM-SF-CCSD[a]	1.17	3.88	4.61

[a]From Ref. [332].

RAS-SF and RAS-SF-*sr*LDAc calculations have been done using the ROHF lowest high-spin triplet [Eq. (4.22)] as the reference configuration, with different RAS2 spaces and the cc-pVTZ basis set. RAS-SF with four electrons in the four frontier π-orbitals in Figure 4.7 produce excitation energies to the E' singlet and triplet states in very good agreement with the EOM-SF-CCSD reference values. On the other hand, transition energy to $^1A'_1$ is largely overestimated (by 1.2 eV). Excitation energies with short-range correlation and the same (4,4) RAS2 space are very close to the wave function results, with slightly underestimated transition to $^1E'$ and too high energy for $^1A'_1$. Hence, we might infer that the short-range DFT correlation approach preserves the wave function accuracy, but *sr*DFTc is not able to recover the missing differential electron correlation in the $^3A'_2 \rightarrow {}^1A'_1$ transition. This correlation is progressively included as we increase the size of the fully correlated space (RAS2). Whereas the transition energy to $^1E'$ and $^3E'$ merely change from the (4,4) to the (8,9) RAS2 spaces, the vertical energy to $^1A'_1$ progressively diminishes by expanding RAS2, with an excitation energy 0.35 eV higher than the EOM-SF-CCSD value in the largest space.

It is also worth noting that the non-iterative RAS-SF-*sr*LDAc0 energies are very close to the self-consistent (RAS-SF-*sr*LDAc) values, as shown for the (4,4) and (8,9) RAS2 spaces.

Overall, these results indicate that, although the *sr*DFTc correction to the RASCI energies might be a good approach for the study of open-shell systems, in some cases it may not be able to capture the delicate correlation balance between different electronic states, such as in the accurate quantification of the excitation energy to the $^1A_1'$ state in TMM.

4.5.6 Gallium dimer

The cohesive energy of metals in group 13 of the periodic table progressively decreases with the atomic number: Al (3.39 eV), Ga (2.81 eV), In (2.52 eV) and Th (1.88 eV). In general, this trend correlates well with their melting point, except for gallium, which exhibits a much lower temperature: Al (934 K), Ga (303 K), In (430 K) and Th (304 K). Such distinctive properties of gallium have been rationalized by the structure of the solid α-Ga phase, which is composed of interacting Ga_2 units, and through the study of small gallium clusters [339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349].

Experimental and theoretical investigations agree with the spin-triplet nature of the Ga_2 ground state ($^3\Pi_u$) in the gas phase [339, 344, 348], with a binding energy experimentally measured around 1.1 eV. Moreover, whereas Ga_2 molecule spontaneously reacts with H_2, atomic gallium requires photo activation in order to interact with dihydrogen to form GaH_2 [350, 351, 352, 353]. This difference in reactivity has been attributed to the wide range of accessible excited states in the dimer [345]. Electronic structure calculations have identified $^3\Sigma_g^-$ as the first excited state being nearly degenerated to $^3\Pi_u$, while singlet excited states, *i.e.*, $^1\Sigma_g^+$, $^1\Pi_u$ and $^1\Delta_g$, have been predicted to lie within a few tenths of an eV above the ground state minimum [343, 345, 347].

Computational characterization of the near-degeneracies between low-lying states of Ga_2 requires the use of methodologies able to account for electron correlations, and represents a challenging test for WFT-*sr*DFT approaches. Here we apply the use of short-range DFT correlation to the computation of ground and excited states of gallium dimer and we compare our results with experimental measurements [339, 344] and previous theoretical calculations [343, 347, 346, 348]. Ground state dissociation energy (D_e) and adiabatic excitation energies to the singlet and triplet electronic states of Ga_2 have been computed using the optimized geometries reported by Himmel and collaborators [345], with bonding distances of 2.763 Å, 2.510 Å, 3.093 Å, 2.781 Å and 2.587 Å respectively for $^3\Pi_u$, $^3\Sigma_g^-$, $^1\Sigma_g^+$, $^1\Pi_u$ and $^1\Delta_g$ states. In all cases we use the $^3\Sigma_g^-$ ROHF determinant as the reference configuration for RAS-SF and RAS-SF-*sr*DFTc calculations with a RAS2 space defined by four electrons in four orbitals (σ_u^+, π_u, and σ_g^+ in Figure 4.9). Previous calculations have suggested that triple ζ basis sets might be too small to describe correlation effects in the gallium dimer and that the accurate prediction of dissociation energy might require the use of quadruple ζ quality basis sets [348, 354]. Thus, to account for basis set dependence effects, we have computed D_e and adiabatic energies with cc-pVXZ (X = D, T and Q) basis.

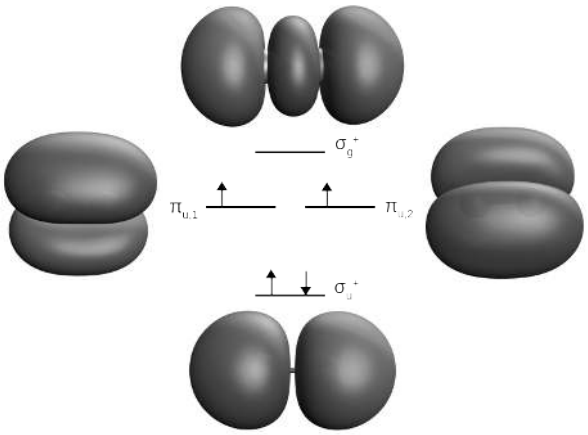


Figure 4.9: ROHF/cc-pVDZ ($^3\Sigma_g^-$) frontier molecular orbitals defining the RAS2 space employed in RAS-SF and RAS-SF-*sr*DFTc calculations of Ga_2.

Dissociation energy of Ga_2 in gas phase has been experimentally determined to be 1.15 ± 0.05 eV [342]. Comparison of this value to computational results need to take into account ZPVE and relativistic corrections. The former has been theoretically evaluated in the order of 0.01 eV [355], while multiconfigurational calculations predict spin-orbit coupling to contribute with 0.10 eV to D_e [356]. Hence, the corrected experimental value of D_e directly comparable with non-relativistic calculations can be taken as 1.26 eV.

RAS-SF with a relatively small fully correlated space (four electrons in four orbitals) is unable to finely recover differential correlation effects between the ground state energy minimum and the dissociation limit, and underestimates D_e by nearly 0.5 eV (Table 4.7). On the other hand, adding short-range DFT correlation in RAS-SF-*sr*DFTc notably increases the computed relative stability of the minimum, with errors below 0.1 eV with respect to the corrected experimental value. It is worth mentioning that the present results do not show a large differences between binding energies computed with the cc-pVTZ and cc-pVQZ basis. Hence, we must conclude that basis sets of triple ζ quality appear to be able to accurately predict the strength of Ga_2 binding, at least within the RAS-SF-*sr*DFTc approach.

Computed adiabatic energies for the low-lying states of Ga_2 are shown in Table 4.7. The triplet-triplet separation energy obtained with RAS-SF and RAS-SF-*sr*DFTc are in good agreement with high-level CCSD(T)/QZVP results [348], specially for RAS-SF-*sr*DFTc with the cc-pVQZ basis. The presence of *sr*DFTc has a much stronger impact on the adiabatic energy of the three excited singlets. Short-range DFT correlation systematically reduces the energy gap to the three excited singlets by 30-40 kJ/mol with respect to RAS-SF energies (and for $^1\Pi_u$ and $^1\Delta_g$ with CASSCF/SVP), which could be related to the lack of dynamic correlations in the wave function methods playing a major role in the stabilization of spin singlet states.

Table 4.7: Binding energies D_e for the Ga_2 $^3\Pi_u$ ground state (in eV) and adiabatic excitation energies (in kJ/mol) computed with RAS-SF and RAS-SF-*sr*DFTc (*sr*DFTc in short) for the low-lying excited states with respect to the ground state.

method	basis	D_e	$^3\Sigma_g^-$	$^1\Sigma_g^+$	$^1\Pi_u$	$^1\Delta_g$
RAS-SF	cc-pVDZ	0.784	0.9	59.2	54.7	69.2
RAS-SF	cc-pVTZ	0.781	1.4	54.1	48.6	62.0
RAS-SF	cc-pVQZ	0.789	1.0	52.8	47.2	59.8
*sr*DFTc	cc-pVDZ	1.174	2.3	26.3	19.0	32.4
*sr*DFTc	cc-pVTZ	1.181	3.6	20.9	14.4	29.8
*sr*DFTc	cc-pVQZ	1.194	4.4	14.7	10.0	28.8
CASSCF[a]	SVP		7.1	19.0	46.3	56.4
CCSD(T)[b]	QZVP		5.0			
Exp.[c]		1.15				
Exp. corr.[d]		1.26				

[a]From Ref. [345]. [b]From Ref. [348]. [c]From Ref. [342]. [d]Experimental value corrected with ZPVE (0.01 eV) from Ref. [355] and relativistic effects (0.10 eV) from Ref. [356].

Finally, we represent the energy profiles for the two triplets and three singlet states of Ga_2 computed at the RAS-SF-*sr*DFTc/6-31G(d) level. The RAS-SF-*sr*DFTc potential energy curves notably resemble previous CCSD(T) results [346], with minimum energy values near the interatomic distances reported by Kohn [345]. The $^3\Sigma_g^-$ state is the most stable one at short bond distances (< 2.5 Å) and with $^1\Delta_g$ they are the two low-lying states with the shortest interatomic separation at the energy minimum (∼2.4 Å). On the other hand, the equilibrium distance is sensibly larger in the $^3\Pi_u$ state (around 2.6 Å) and even larger for the other two excited singlets ($^1\Sigma_g^+$ and $^1\Pi_u$).

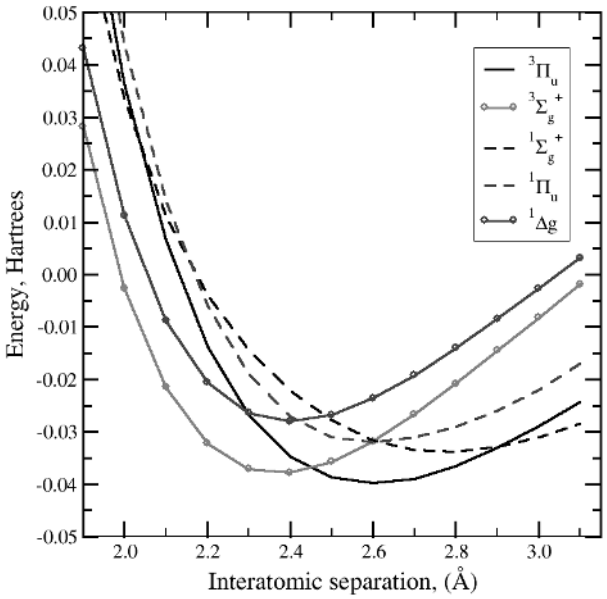


Figure 4.10: Dissociation energy curves (in Hartrees) of the ground and low-lying states of the Ga_2 computed at the RAS-SF-*sr*DFTc/6-31G(d) level. Energies are given with respect to the dissociation limit.

4.6 Conclusions

In this Chapter we have explored the suitability of WFT-*sr*DFT to characterize electronic states of atomic and molecular systems with strong open-shell character. Our results have revealed the limitations of WFT-*sr*DFT to provide a balanced description between closed and open-shell configurations. Detailed analysis of a two-electron system, *i.e.*, the H_2 dissociation, has helped to identify the inability of short-range DFT to incorporate spin polarization effects in open-shell systems as described by spin complete wave functions. Moreover, these issues, emerging from the mingle of WFT and DFT through the range-separation of the Coulomb operator, are especially important in the (short-range) DFT exchange energy functional. These observations have motivated us to develop different variants of the WFT-*sr*DFT scheme based on strategies to include spin polarization effects into the DFT part (*sr-p*DFT), as already employed in other WFT-DFT hybrid methodologies, and to modify the short-range exchange contribution. For the latter, we have consider two options: (i) short-range HF exchange (*sr*-EXX), and (ii) exact exchange as described by the multiconfigurational wave function (WFT-*sr*DFTc).

The performance of these different WFT-*sr*DFT flavors has been tested in the dissociation of the H_2 molecule, the evaluation of energy gaps in triplet atoms and organic diradical molecules, and to the characterization of the low-lying electronic states of the gallium dimer. The spin polarization scheme shows notable improvements with respect to WFT-*sr*DFT, but this is in essence a state specific method that should not be combined with the state average approach, which complicates its use in the computation of excited states. The replacement of the short-range (pure) DFT exchange for HF exchange does not improve the results, which actually become much worst. Hence *sr*-EXX or the use of short-range hybrid exchange seems to be a path not worth pursuing. Finally, the re-integration of the short-range exchange into the wave function part (WFT-*sr*DFTc) notably and systematically improves the WFT-*sr*DFT and pristine WFT results. Comparison of the self-consistent inclusion of short-range DFT correlation to its one-shot approximation shows only small numeric differences, and in some cases the non-iterative approach even improves the computed values, since it does not require the use of the state average approximation in the simultaneous computation of multiple roots of the electronic Hamiltonian.

In spite of the encouraging results of WFT-*sr*DFTc (and WFT-*sr*DFTc0) obtained in the present study, it is necessary to further evaluate their accuracy in the computation of ground and excited state relative energies in a wider range of molecular systems.

CHAPTER 5

Small-occupation density functional correlation energy correction to wave function approximations

5.1 Abstract

In this work, we introduce a novel hybrid approach, termed WFT-*so*DFT, designed to seamlessly incorporate DFT correlation into wave function ansatzes. This is achieved through a partitioning of the orbital space, distinguishing between large and small natural occupation numbers (NOONs) associated with wave function theory (WFT) and DFT correlation, respectively. The method uses a novel criterion for partitioning the orbital space and mapping the electron density in natural orbitals with small occupation with the correlation energy of fast electrons within the homogeneous electron gas. Central to our approach is the introduction of a separation parameter ν, the choice of the WFT approach, and the correlation functional. Here, we combine the RASCI wave function with *hole* and *particle* truncation with a local density correlation functional to only account for small-occupation correlation energy. We investigate the performance of the method in the study of small but challenging chemical systems, for which WFT-*so*DFT demonstrates notable improvements over pristine wave function calculations. These findings collectively highlight the potential of the WFT-*so*DFT approach as a computationally affordable strategy to improve the accuracy of WFT electronic structure calculations.

5.2 Introduction

Molecular electronic structure descriptions traditionally rely on the non-relativistic framework and the Born-Oppenheimer approximation, simplifying the treatment of nuclei and electrons to solving the electronic time-independent Schrödinger equation. Although the full configuration interaction (FCI) offers an exact solution (in a given one-electron basis), its computational demands limit its applicability to small systems [5, 6, 7]. These limitations have spurred the development of approximate quantum chemistry methods, with a focus on reducing the dimensionality of the problem in two main directions: (i) employing finite basis sets for the description of single electrons, which already implies a level of approximation

with respect to the exact solution [357], and (ii) addressing electron-electron interactions, also known as the electron correlation problem [15]. Amongst the latter, Hartree-Fock (HF) theory provides a qualitative solution for electronic ground states [358], capturing a large portion of the total electronic energy in closed-shell molecules. Post-HF methods build on this foundation, employing perturbation theory (e.g., MP2), configuration interaction (CI), or coupled cluster (CC) methods (as in CCSD[67] or CCSD(T)[37]). However, HF theory struggles in strongly correlated systems, necessitating multiconfigurational approaches like the complete active space self-consistent field (CASSCF)[58] or its perturbative variant (CASPT2),[59] which are computationally intensive and thus restricted to smaller molecules. On the other hand, density functional theory (DFT), in particular in the Kohn-Sham scheme [359], offers an efficient alternative by bypassing the need for many-body wave functions. However, the lack of an exact universal energy functional limits the accuracy of DFT approximations [68], especially in cases of degeneracies or near-degeneracies.

Given the distinct capabilities of wave function theory (WFT) and DFT in capturing static and dynamic correlations respectively, there's been interest in combining these approaches [73]. Common mingled WFT-DFT models [360] rely on range-separation of the Coulomb operator [77, 315], on the multiconfiguration pair-density functional theory (MC-PDFT) [183, 187], or the multiconfiguration density coherence functional theory (MC-DCFT) [361], showing promise in various chemical scenarios. Alternatively, it has been shown that the eigenstates of the density matrix, i.e., natural orbitals (NOs), and their occupation numbers (NOONs) can be used to characterize correlation effects [222]. Analysis of NOONs has inspired novel strategies for active space selection [362, 155, 363, 364], and the distribution of NOONs has been used to characterize strongly correlated systems [223] and as a metric to quantify the open-shell character of electronic wave functions [365, 241, 366, 367]. Results from condensed matter physics, where strongly correlated electrons are usually described with model Hamiltonians [21], and the homogeneous electron gas (HEG), where the NOs are plane waves [222], show that large NOONs describe molecule-specific effects, while those orbitals with small occupations and large momentum can be related to short-range effects. These properties were exploited by Savin, who designed a strategy to mix WFT and DFT by splitting the orbital space between those NOs with NOONs larger than a given threshold (evaluated with WFT), and those with small NOONs (used to quantify electron correlation with DFT) [22, 22, 23, 172, 368]. But, since those initial works, no further investigations have been carried out in this direction. In the present work, we reformulate the correction of multiconfigurational wave functions with DFT correlation through orbital space separation. We aim to establish a general formalism to accommodate for suitable WFT models and DFT functionals based on the splitting of the orbital space, termed WFT with small-occupation DFT (WFT-*so*DFT), and evaluate its performance, addressing its advantages and limitations in characterizing chemical systems.

This chapter is organized as follows. First, we introduce and describe the theoretical framework for the mingling of WFT with DFT (section 5.3), we motivate the partition of the orbital space as a justified strategy and we discuss several aspects of the methodology (section 5.4). Section 5.5 describes the workflow of the WFT-*so*DFT method, and the employed computational details and nomenclature. Section 5.6 addresses the performance of the computation of the He and Be atomic series, the single bond breaking in H_2 and C_2H_4, and multiple bond dissociation in N_2 dissociation. The main findings of the work are summarized in the conclusions section.

5.3 General framework to merge WFT with DFT

In WFT, the exact (non-relativistic) ground state energy for an electronic molecular system can be obtained variationally as:

$$\mathcal{E}_0 \equiv E_0[\Psi_0] = \min_{\Psi} \Big\{ \langle \Psi | \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} | \Psi \rangle \Big\}, \tag{5.1}$$

where Ψ_0 is the exact ground state wave function, $\hat{T}$ is the kinetic energy operator, $\hat{V}_{ne}$ and $\hat{V}_{ee}$ the operators respectively accounting for nuclei-electron and electron-electron interactions, and the minimization runs over all possible antisymmetric wave functions, i.e., the entire Hilbert space. In practice, the minimization in Eq. 5.1 is conducted within a finite Hilbert subspace (S) defined by the electron correlation approximation and the chosen basis set,

$$\mathcal{E}_0^{\text{WFA}} = \min_{\widetilde{\Psi} \in S} \Big\{ \Big\langle \widetilde{\Psi} \Big| \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} \Big| \widetilde{\Psi} \Big\rangle \Big\}. \tag{5.2}$$

The fairness of the wave function approximation (WFA) is given by the difference between the exact and approximate energies. In particular, for mean-field wave functions, i.e., Hartree-Fock (HF), the difference with respect to the exact energy is called *electron correlation energy*. More concretely, for wave functions capturing the main electronic configurations, e.g., HF (only one) or CASSCF [58] with small or moderate active spaces, the energy error of the approximate wave function is labelled as *dynamic correlation energy* [43, 46].

Alternatively, the Hohenberg-Kohn theorem [8] allows to express the ground state energy as a function of the electron density,

$$E_0[\rho] = F_0[\rho] + \int v(r)\rho(r)dr, \tag{5.3}$$

where $v(r)$ is the external potential, $v(r) = \langle \Psi | \hat{V}_{ne} | \Psi \rangle$, and $F_0[\rho]$ is the universal Levy-Lieb (LL) density functional [194, 369],

$$F_0[\rho] = \min_{\Psi \to \rho} \Big\{ \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle \Big\}, \tag{5.4}$$

with the minimization performed over all wave functions with density ρ. Then, the exact ground state energy can be obtained by minimizing the energy functional in Eq. 5.3,

$$\mathcal{E}_0 \equiv E_0[\rho_0] = \min_{\rho} \Big\{ F_0[\rho] + \int v(r)\rho(r)dr \Big\}, \tag{5.5}$$

where ρ_0 is the exact ground state density, $\rho_0 \equiv \rho[\Psi_0]$. If the LL functional is obtained only considering a restricted wave function space, $\widetilde{\Psi} \in S$, and assuming that the restricted space S contains at least one wave function $\widetilde{\Psi}$ able to generate ρ,

$$\widetilde{F}_0[\rho] = \min_{\widetilde{\Psi} \to \rho} \Big\{ \Big\langle \widetilde{\Psi} \Big| \hat{T} + \hat{V}_{ee} \Big| \widetilde{\Psi} \Big\rangle \Big\}, \tag{5.6}$$

then, it is possible to define a correlation energy functional, $E_c[\rho]$, accounting for the missing energy to the exact ground state,

$$E_c[\rho] = E_0[\rho] - \widetilde{E}_0[\rho] = F_0[\rho] - \widetilde{F}_0[\rho]. \tag{5.7}$$

Hence, the exact ground state energy can be formally obtained by computing the approximate and correlation functionals at ρ_0. In practice, though, one does not have access to ρ_0.

Therefore, we approximate the exact energy by evaluating each term at the approximate density, that is, the optimal density within the restricted space, $\widetilde{\rho}_0 \equiv \rho[\widetilde{\Psi}_0]$:

$$\mathcal{E}_0 \approx \widetilde{E}_0[\widetilde{\rho}_0] + E_c[\widetilde{\rho}_0]. \tag{5.8}$$

Eq. 5.8 can be used to define new methods mixing approximate wave functions with DFT by assigning the calculation of the approximate ground state energy to WFA (Eq. 5.2):

$$\mathcal{E}_0 \simeq \mathcal{E}_0^{\text{WFA}} + E_c[\widetilde{\rho}_0]. \tag{5.9}$$

In general, for a good enough trial wave function $\widetilde{\Psi}_0$, one should expect that the errors derived from the explicit form of $E_c[\rho]$ to be larger than those derived by the use of approximate densities, thus justifying the use of Eq. 5.8.

We must stress that this strategy relies on the proper separation of the WFT and DFT terms in order to compute complementary contributions to the electronic energy while avoiding double counting of correlation effects. In the present work, we explore the use of the orbital space as to split the total electronic energy in WFT and DFT contributions.

5.4 Orbital separation based on electron occupation

Eq. 5.9 provides a general framework for the mixing of WFT and DFT. However, it is nowhere close to being a working equation. Reliable choices for explicit forms of the two contributions in the LHS of Eq. 5.9, i.e., approaching to the exact energy ($\mathcal{E}_0$), should consider the properties and limitations of the available WFAs and DFAs. Hence, we envision the use WFA to account for strong (non dynamic) correlation effects (difficult to recover with energy density functionals) [159, 370], leaving dynamic correlation to DFA, for which WFAs typically show slow convergence.

The direct additive combination of WFA energy with the correlation energy from DFT, although tempting, it has been shown to be inappropriate, since the DFT correlation contribution might include correlation effects already present in the wave function part, and vice-versa (double counting of electron correlations) [371]. Therefore, it is necessary to employ schemes for the proper separation of the two contributions. To that aim, we rely on the momentum distribution ($n(k)$) of the homogeneous electron gas (HEG), with a general profile shown in Figure 5.1, corresponding to a distribution of the electron occupation expressed in terms of the NOs of the system (plane waves). The probability of finding an electron with $k < k_F$ is close to one up until the electron momentum approaches the Fermi wavevector (k_F), where it exhibits a small decay with a profile depending on the electron density [372], as characterized by the Wigner-Seitz radius (r_s). The electron density of the HEG presents a discontinuity at k_F. For the case of interacting electrons, the momentum distribution $n^+ \equiv n(k \to k_F^+, r_s) \neq 0$ (+ superscript indicates right-handed limit), whereas $n(k)$ rapidly decays as $k \to \infty$.

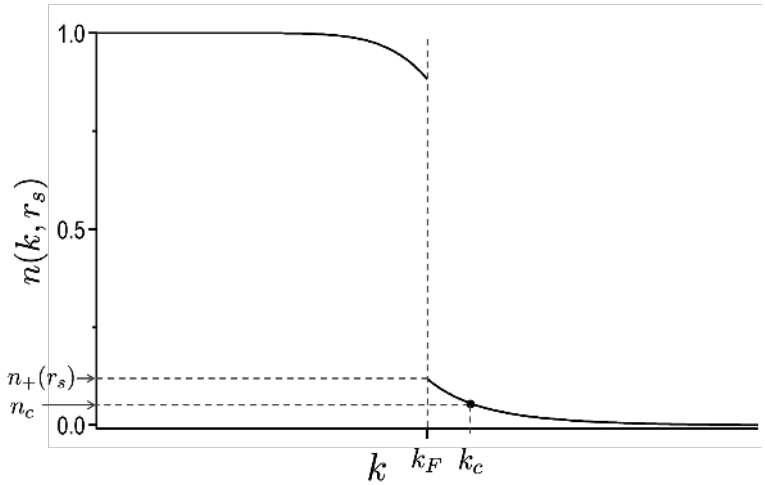


Figure 5.1: Representation of the momentum distribution for the HEG. Two points are highlighted, the first at the Fermi wavevector (k_F) with its associated probability limit approaching from $k > k_F$ ($n_+(r_s) = n(k_F^+, r_s)$) and the second with a momentum $k_C > k_F$ and probability n_C that corresponds to the point from which the dynamic correlation can be accounted by the DF.

It has been argued that the dynamic correlation in the HEG can be associated with the high kinetic energy electrons, that is, to those NOs with small occupancies [222]. This idea suggests a NO-based criterion to split between non-dynamic and dynamic correlations, which can be respectively addressed by WFA and DFA. Savin connected such behavior of the HEG to the study of atomic and molecular systems by suggesting to compute the contribution to the electron energy of NOs with high NOONs with a WFA able to recover the system-specific electron correlation effects and evaluate the (dynamic) correlation energy of NOs with small NOONs with DFT correlation energy functionals [22, 22, 23]. Accordingly, it is possible to define a WFT-DFT composite approach, that we call *small-occupation DFT corrected WFT* (WFT-*so*DFT), by applying the occupation-based orbital splitting to Eq. 5.9,

$$E^{\text{WFT-}so\text{DFT}} = \mathcal{E}_0^{\text{WFA},\nu} + E_c^{so,\nu}[\rho], \tag{5.10}$$

$$\mathcal{E}_0^{\text{WFA},\nu} = \min_{\Psi^\nu} \langle \Psi^\nu | \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} | \Psi^\nu \rangle, \tag{5.11}$$

where ν is the (positive-definite) NOON separation parameter, i.e., the molecular counterpart of n_C for the HEG in Fig. 5.1, Ψ^ν is the wave function obtained by the chosen WFA, but only considering NOs with NOONs $> \nu$. Superindex ν explicitly indicates the dependence of the wave function and the correlation functional on the ν-parameter. Three main issues need to be considered in the use of Eq. 5.10: (i) the choice of the occupation-separation parameter ν; (ii) the design of correlation energy functionals (only) corresponding to NOs with low NOONs; and (iii) the accurate description of the NOONs distribution of the system. These three key aspects of WFT-*so*DFT will be discussed in the following subsections.

5.4.1 Occupation separation parameter

The choice of the occupation threshold ν splitting the orbital space in WFT-*so*DFT is controlled by the ν-parameter. It is important to notice that Eq. 5.10 represents a correction to the WFA. In the limit of $\nu \to 0$ the method converges to the WFT, while the $\nu \to \infty$ limit provides the total DFT electron correlation energy.

In the original formulation [22, 23], the choice of ν was related to the idea of the existence of a universal threshold value ν_c, which in practice can be used to define the system-dependent ν as:

$$\nu = \max_{n_i}\{n_i \leq \nu_c\}. \tag{5.12}$$

Numerical analysis on neutral systems recommended the use of $\nu_c \approx 0.01$ a.u. [23]. Such approach, although simple, only considers the NOON occupation of a single NO, and completely disregards the rest of small-occupation NOs. To overcome these potential limitations, we introduce a new definition of ν by considering the entire tail of small NOONs. For that, we define ν through Eq. 5.13,

$$\int_{k_c}^{\infty} n(k, r_s)dk = \frac{1}{N}\sum_{i}^{n_i \leq \nu_c} n_i \tag{5.13}$$

where

$$\int_{0}^{\infty} n(k, r_s)dk^3 = 1, \tag{5.14}$$

k is in units of k_F, $\{n_i\}$ are the NOONs obtained from the diagonalization of the one-particle density matrix (1-PDM), and N is the total number of electrons. Compared to Eq. 5.12, we envision Eq. 5.13 to be more robust to the specificities of the discrete distribution of NOONs as obtained with WFAs, e.g., NOONs degeneracies around ν_c, and that it might represent a more balanced manner to include *so*DFT correlation energy. Then, ν can be obtained either as $\nu = n(k_c, r_s)$ (defining the ν_c threshold and computing k_c with Eq. 5.13), or alternatively as $\nu = \nu_c$ (defining k_c and computing ν_c from Eq. 5.13) [373].

The use of Eq. 5.13 requires to have an explicit expression for $n(k, r_s)$ distribution within the $k_c \leq k \leq \infty$ range. Gori-Giorgi and Ziesche[374] have shown that $n(k, r_s)$ for $k > k_F$ can be approximated as

$$n(k > k_F, r_s) \approx n_+(r_s)\frac{n(k \to \infty, r_s)}{n(k \to \infty, r_s)|_{k=k_F}} \tag{5.15}$$

where $n_+(r_s) = n(k_F^+, r_s)$ (Figure 5.1). The momentum distributions at the $k \to \infty$ limit is known to decay proportionally to k^{-8} [375, 376],

$$n(k \to \infty, r_s) = \frac{8(\alpha_0 r_s)^2 g_0(r_s)}{9\pi k^8} + \mathcal{O}(\frac{1}{k^{10}}), \tag{5.16}$$

where α_0 is the Bohr radius and $g_0(r_s)$ is the value of the pair-distribution function at zero-interelectronic distance, which can be parametrized as[377]

$$g_0(r_s) = \frac{1 - Br_s + Cr_s^2 + Dr_s^3 + Er_s^4}{2}e^{-dr_s}, \tag{5.17}$$

with parameters $d = 0.7524$, $B = 0.7317 - d$, $C = 0.08193$, $D = -0.01277$, and $E = 0.001859$. At this point, only a suitable expression for $n_+(r_s)$ remains to be defined. Here, we will consider the form proposed by Savin [22],

$$n_+(r_s) = \frac{r_s}{r_s + 8.45}, \tag{5.18}$$

and the one given by Gori-Giorgi and Ziesche [374],

$$n_+(r_s) = \frac{q_1 r_s}{1 + q_2 r_s^{1/2} + q_3 r_s^{7/4}}, \tag{5.19}$$

where $q_1 = 0.088519$ (from RPA), $q_2 = 0.45$, and $q_3 = 0.022786335$.

It is worth noticing that the orbital separation strategy discussed here might not preserve additive separability. Evaluation of size consistency errors in the WFT-*so*DFT scheme can be found in the Supporting Information.

5.4.2 Small-occupation correlation energy functionals

The challenging aspect in the design of small-occupation correlation energy functionals ($E_c^{so,\nu}[\rho]$) is related with the need to exclude the contribution of NOs with large NOONs from the DFT correlation. To that end, Savin proposed a ν-dependent correlation functional that within the local density approximation (LDA) takes the form

$$E_c^{so,\nu}[\rho] = \int \rho\epsilon_c(\rho)\varphi(\rho,\nu)dr. \tag{5.20}$$

The function $\varphi(\rho,\nu)$ introducing the dependence on ν was adjusted to the analytical form

$$\varphi(\rho,\nu) \approx \left(\frac{\nu}{n_+(r_s)}\right)^{\gamma} \tag{5.21}$$

for $\nu < n_+(r_s)$, where the γ value was fitted to 0.329 by using the momentum distribution from Pajanne and Arponen [378] and coupled-cluster calculations from Freeman [379]. For $\nu > n_+(r_s)$, $\varphi(\rho,\nu)$ can be fixed to 1 by imposing $\epsilon_c(\rho)\varphi(\rho,\nu) \leq \epsilon_c(\rho)$, $\forall r$. Hence, for large ν Eq. 5.20 recovers the full DFT correlation, whereas for $\nu \to 0$ and the DFT correlation energy vanishes.

5.4.3 Sampling the orbital occupation space

The WFT-*so*DFT strategy relies on having access to an accurate profile of the NOONs of the system. Therefore, the WFA should be able to produce a distribution close to the exact one, in particular in the small-occupation region, which is the one eventually dictating the *so*DFT correlation energy through the ν-parameter. Hence, the use of compact multiconfigurational approaches (featuring a very small set of configurations) to generate NOs and their NOONs, although they could provide a good approximation of ρ, they might be insufficient for the accurate assessment of $E_c^{so,\nu}[\rho]$. Similarly, small basis sets might not provide the flexibility necessary to characterize $n(k > k_F, r_s)$. On the other hand, employing highly correlated wave function approaches, e.g., CASSCF with a large active space, would inevitably increase the computational demands of the method and defeat the purpose of the WFT-*so*DFT strategy. Therefore, in principle, an ideal ansatz should present two important features: (i) a small number of configurations capable to recover the system-dependent features (strong correlations), and (ii) a limited set of additional terms able to describe the distribution of small-occupation NOONs.

Here we propose the restricted active space CI (RASCI)[6] method within the one *hole* and one *particle* approximation as implemented in the Q-Chem package [380, 115].

Recalling that, the single reference RASCI family of methods emerges from splitting the orbital space of a reference configuration (ϕ_0) into three different subspaces: RAS1, RAS2 and RAS3, with RAS1 (RAS3) orbitals being doubly occupied (virtual) in ϕ_0. In order to achieve the two indicated requirements of WFAs for the generation of the NOON distribution, we consider RASCI wave functions with a small RAS2 space (to capture strong correlations) and

the truncation of the excitation operator to $\hat{R} = \hat{r}_0 + \hat{r}_h + \hat{r}_p$, that is, the *hole* and *particle* approach (RASCI(h, p)), as schematically shown in Figure 2.1.

To exemplify how the chosen ansatz is able to evaluate the NOONs distribution, we compute the ground state NOs and NOONs for the all-*trans* octatetraene (C_8H_{10}) with different truncations of $\hat{R}$. In this case, pristine RASCI wave functions (with no soDFT contribution) were obtained by applying a quadruple spin-flip operator to the high-spin ($S = 4$) ROHF reference. The RAS2 space was defined by the 8 electrons in the 8 singly-occupied π-orbitals in the ROHF reference, with RAS1 and RAS3 including all doubly occupied and doubly unoccupied orbitals, respectively. The molecular geometry was optimized at the MP2/6-311(2+,2+)G(d,p) computational level [125]. Similar qualitative results were obtained for other molecules (see Appendix B).

Figure 5.2a shows the profiles of the low-occupation tail of ground state NOONs computed with $\hat{R} = \hat{r}_0$ (RASCI(0)), $\hat{R} = \hat{r}_0 + \hat{r}_h$ (RASCI(h)), $\hat{R} = \hat{r}_0 + \hat{r}_p$ (RASCI(p)), and $\hat{R} = \hat{r}_0 + \hat{r}_h + \hat{r}_p$ (RASCI(h, p)). The RASCI(0) truncation, i.e., CASCI, is unable to populate the orbital space beyond RAS2. It produces a set of NOs (with the dimension of RAS3) with vanishing NOONs. Hence, this approach, like CASSCF with small and moderate active spaces, cannot be used as a WFA in WFT-soDFT. Including $hole$ excitations does not provide any change in the small-occupation region. This is expected, since RASCI(h) only allows orbital mixing between RAS1 and RAS2 spaces, and can only generate fractional NOONs in the large-occupation region and the strongly correlated space (RAS2). On the other hand, allowing excitations into the (much larger) RAS3 space, RASCI(p), produces a finite distribution of NOs with small NOONs. Therefore, we conclude that the presence of $particle$ terms is mandatory. In this example, allowing both type of excitations does not involve any appreciable change in the description of $n_+(k > k_F, r_s)$ with respect to RASCI(p), but the simultaneous presence of both type of excitation allows the mixing of the entire orbital space (RAS1+RAS2+RAS3). Moreover, the computational cost associated to $hole$ excitations is, in general, considerably smaller than that related to $particle$ configurations. In other words, the size of RAS3 (proportional to the number of basis set functions) is much larger than RAS1 (proportional to the system size). Therefore, RASCI(h, p) will be the WFA used along the rest of the study.

The importance of $particle$ excitations in the sampling of $n_+(k > k_F, r_s)$ is directly related with the basis set employed, as it defines the size and properties of the RAS3 space. Figure 5.2b shows the small-occupation profiles for the ground state of octatetraene obtained with RASCI(h, p) and with the cc-pVXZ (X = D, T, and Q) series of basis. Increasing the size of the basis set modifies the profile of the NOON distribution, with an overall electron occupation of the NOs with the smallest occupations (integration of the NOONs of NOs beyond RAS2 space) increasing with the dimension of the basis set: 8.3×10^{-3} (cc-pVDZ), 9.8×10^{-3} (cc-pVTZ), and 11.1×10^{-3} (cc-pVQZ) electrons. Therefore, as desired, increasing the size of the basis set will trigger a larger soDFT correlation correction energy into WFA.

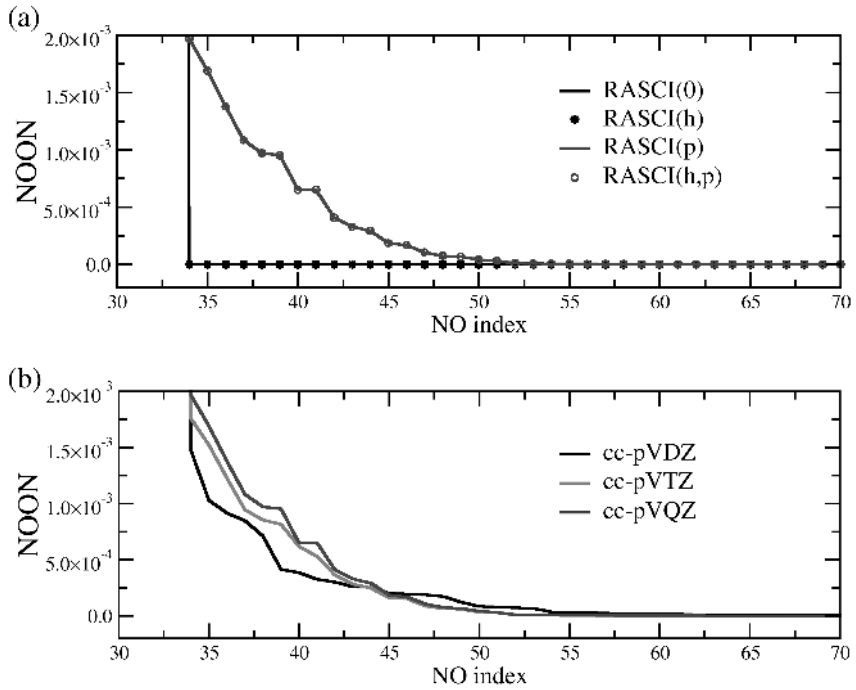


Figure 5.2: Ground state NOONs at the tail of the NO space of octatetraene computed with: (a) the cc-pVQZ basis and different truncations of $\hat{R}$, and (b) RASCI(h, p)/cc-pVXZ with X = D, T and Q.

5.5 Methods

5.5.1 Computational Workflow

The general procedure for the computation of the WFT-*so*DFT energy can be schematically described as follows:

1. Initial WFA calculation to obtain NO and NOON pairs (Eq. 5.2).
2. Select occupation-separation parameter ν (Eq. 5.12 or Eq. 5.13).
3. Compute *so*DFT correlation energy: $E_c^{so,\nu}[\rho]$ (Eq. 5.20).
4. Compute the ν-dependent WFA energy: $\mathcal{E}_0^{\text{WFA},\nu} = \min_{\Psi^\nu} \langle \Psi^\nu | \hat{H} | \Psi^\nu \rangle$ (Eq. 5.11).
5. Obtain the total energy by adding both contributions (Eq. 5.10).

Note that the *so*DFT correlation energy (step 3) is computed employing the electron density obtained from the initial WFA calculation (step 1). This is justified by the fact that, in general, the initial WFA is designed to provide good NO and NOON pairs, i.e., with an electron density expected to mildly deviate from the exact one (assumption in Eq. 5.8). The ν-dependent WFA energy (step 4) is computed with the chosen WFA, but disregarding the NOs with NOONs smaller than the ν-parameter.

5.5.2 Computational details and nomenclature

All WFT calculations have been done with the RASCI(h, p) methodology with the RAS2 space specifically selected for each case and with RAS1 (RAS3) including all orbitals doubly occupied (unoccupied) in the reference configuration (not included in RAS2). The nature of the excitation operator $\hat{R}$ in RASCI(h, p) [125, 126]: excitation energy (EE), spin-flip (SF), ionization potential (IP) or electron attachment (EA), is specified in each example (see section 5.6). Computation of Ψ^{ν} in step 4 was obtained by removing those NOs with NOONs smaller than ν. Unless indicated, the dimension of RAS2 used to obtain $E^{\mathrm{WFA},\nu}$ was the same than the one employed in the initial WFA calculation. In practice, the computation of the DFT correlation energy has to rely on an approximation, i.e. density function approximation (DFA). Here we employ the Vosko, Wilk, and Nusair (VWN) LDA correlation functional [305]. Possible extensions to other LDA or generalized gradient approximation (GGA) correlation functionals, or the use of fictitious spin density to mimic the effects of unpaired electrons (as in spin-density functionals) have not been explored in the present work. The introduced methodology, with RASCI(h, p) as the WFA and soVWN as the small-occupation correlation energy functional, will be labelled as RAS-soVWN. When necessary, the S or G labels will be used to explicitly indicate the Savin's (RAS-soVWN(S), Eq.5.18) or Gori-Giorgi's (RAS-soVWN(G), Eq.5.19) form of $n_+(r_s)$, respectively. In order to evaluate the impact of the (potential) double counting of correlation effects, for some of the systems studied in section 5.6 we explore the total electronic energy obtained by the simple addition of the soDFT correlation energy to the pristine WFA energy (Eq. 5.2), that is, skipping step 4 in the computational workflow. These calculations can be found in the Supporting Information. Unless indicated, the selection of the occupation-separation parameter ν has been done by taking into account all small-occupation NOONs (Eq. 5.13) by setting a threshold to the integral on the LHS of Eq. 5.13 (equivalent to fixing k_c). Results obtained with the use of a single NOON instead (Eq. 5.12) are shown in the Supporting Information. The computation of $\mathcal{E}_0^{\mathrm{WFA},\nu}$ (Eq. 5.11) is performed at the RASCI level in the NO basis with the same RAS1 and RAS2 dimensions set for the initial RASCI calculation of NOs and their NOONs. In order to ensure that strong correlations are dealt by WFT, the minimum orbital dimension included in the WFA contribution, i.e., the number of NOs, is never smaller than RAS1+RAS2 space (even if the ν-threshold has not been reached). Analysis of the dependence of the soDFT energy with respect to parameter ν can be found in the Supporting Information.

Configuration interaction using a perturbative selection done iteratively (CIPSI) calculations have been performed with the Quantum package by setting a maximum of two million of determinants entering in the internal space, and PT2 energy threshold below 10^{-6} a.u. [113]. All other electronic structure calculations have been carried out with the Q-Chem package [381]. The WFT-soDFT approach has been implemented in combination with RASCI wave functions in a development version of Q-Chem.

5.6 Results and Discussion

5.6.1 Helium and beryllium atomic series

The He and Be isoelectronic series, consisting of atoms and ions with the same electron count but varying nuclear charges (Z), play a pivotal role in evaluating electronic structure methods and have been widely used to calibrate dynamic and nondynamic correlation effects [49, 47]. As Z increases, dimensional arguments suggest that the radius of the electron cloud decreases as Z^{-1}, the electron kinetic energy increases as Z^2, the electron-nuclear potential energy is proportional to $-Z^2$ and the electron-electron potential energy (the Hartree part) is linear with Z [382]. Therefore, the one-body part of the Hamiltonian dominates over the electron-electron interaction, and the correlation energy should be amenable to a perturbative treatment.

In the He series, the ground state remains energetically isolated up to large Z, suggesting that dynamic correlation effects are far more important than the nondynamic correlations (somehow related to degeneracies). Perturbation theory indicates that correlation energy converges to a constant value at the $Z \to \infty$ limit. In this case, LDA overestimates correlation energies that increase with Z, yielding 3.0 eV for He and 5.5 eV for Ne^{8+} [23], while the anticipated values should remain around 1 eV for the entire series [383, 23]. Electron density increases with Z, as a consequence, for larger values of Z, LDA samples the correlation energy of very dense ($r_s \to 0$) electron gases, where the continuous energy level spacing and long-ranged Coulomb interaction trigger a $\ln r_s$ divergence of the energy density [384]. On the other hand, in the (neutral) Be atom the 2s and 2p orbitals are nearly degenerate (quite small HOMO-LUMO gap). As Z increases with the number of electrons fixed to 4, the orbital gap increases but the relative gap (the gap divided by the average energy of the orbitals) actually decreases (type B of nondynamic correlation) [244]. For $Z \to \infty$, E_c diverges linearly. This feature is not reproduced by standard correlation density functionals, e.g., $E_c^{LDA} \sim \ln Z$ [385]. In the exact perturbative treatment, the energy gap diminishes swiftly, with the spacing of energy levels in the denominator playing a pivotal role in determining the correlation energy.

In order to check the performance of our RAS-soDFT approach, we compute the electronic energies for these two atomic series. Figure 5.3 shows the errors of RASCI(h, p) and RAS-soVWN with respect to FCI energies [386]. All calculations have been done with a RAS2 space including 5 orbitals and all electrons (2 in the He series and 4 in the Be series), and the cc-pVQZ basis set.

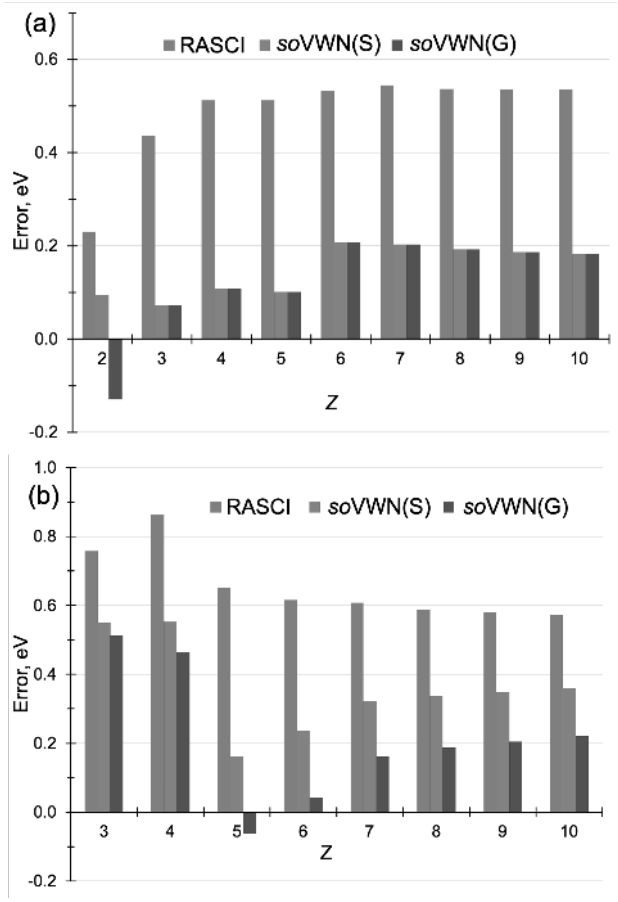


Figure 5.3: RASCI(h, p) and RAS-soVWN (labelled as soVWN with $\nu = 0.001$) energy errors (in eV) with respect to FCI values computed with the cc-pVQZ basis set for the He (a) and Be (b) atomic series. S and G in parenthesis indicate the two considered expressions for $n_+(r_s)$, Eqs. 5.18 and 5.19, respectively.

Despite their simplicity as two- and four-electron systems, both series present significant challenges for RASCI(h, p). Even with the inclusion of five orbitals in RAS2 (1s, 2s, and 2p orbitals), the discrepancies relative to FCI are substantial, on the order of 0.5 eV (approximately 0.2 eV for He) along the He series (see Figure 5.3a) and even larger (0.6–0.9 eV) in the Be series (see Figure 5.3b). RAS-soVWN consistently and significantly mitigates these discrepancies. Remarkably, within the two-electron series, with the exception of the He atom, the electron density predominantly resides within the first five NOs, resulting in almost negligible soDFT correlation energies (a vanishing tail of small occupation NOONs as shown in Figure 3). As a consequence, the enhancement in the correlation energies is directly linked to $\mathcal{E}_0^{\mathrm{WFA},\nu}$, which reflects the utilization of NOs in the variational optimization of the ground state wave function. This pattern holds true except for the He atom, where a substantial contribution from DFT correlation correction is observed. This contribution becomes more pronounced when employing the $n_+(r_s)$ function proposed by Gori (Eq. 5.19). The scenario is quite distinct in the Be series (Figure 4), where there is little to no enhancement observed in $\mathcal{E}_0^{\mathrm{WFA},\nu}$ compared to RASCI (Figure 6). However, a substantial correction arises from soDFT correlation in the atomic energies, with notably significant effects for $Z \geq 5$.

5.6.2 H_2 dissociation

A bond breaking/formation process can be described as a continuous closed- to open-shell transition. The σ-bond dissociation in H_2 is the simplest and most paradigmatic example of such transition. As the molecule is stretched, the gap between the σ and σ^* orbitals, HOMO and LUMO, decreases, facilitating the electron population of the σ^* orbital. As a consequence, the singlet ground state wave function is no longer well described with a single configuration [387, 388]. Hence, a method able to describe the H_2 PES in a balance manner, should treat dynamic correlation between the two electrons (important at equilibrium) and nondynamic correlation (crucial at long distances).

Here, we assess the performance of our *so*DFT correction on RAS-SF/cc-pVDZ energies computed for the H_2 dissociation process. We utilize the ROHF triplet as the reference configuration and a RAS2 space with 2 electrons in 2 orbitals. The pristine wave function solution yields a smooth energy profile, converging to the exact (FCI) solution at large inter-atomic separations, albeit with a slight overestimation of the electronic energy at equilibrium (see Figure 5.4a). Therefore, our aim is to have the *so*DFT approach recover the missing correlation energy at short distances while approaching zero at large distances. Indeed, the NOONs at equilibrium display a tail of small occupancies, contributing to a *so*DFT energy that leads to a lower overall RAS-*so*VWN energy (closer to the FCI solution). Conversely, as the bond distance increases, the spectral resolution of the density matrix evolves towards two singly occupied NOs (with no small occupation tail), i.e., $E_c^{so,\nu} \to 0$ (see Figures 5.4b and 5.4c). Therefore, overall RAS-*so*VWN improves the RAS-SF results by incorporating electron correlation effects at short distances. We note that at very short inter-atomic separation (< 0.7 Å) the *so*DFT correlation ill-behaves (not shown in Figure 5.4c). Complete profiles for the sum of NOONs and $E_c^{so,\nu}$ can be found in Appendix B (Figure 7).

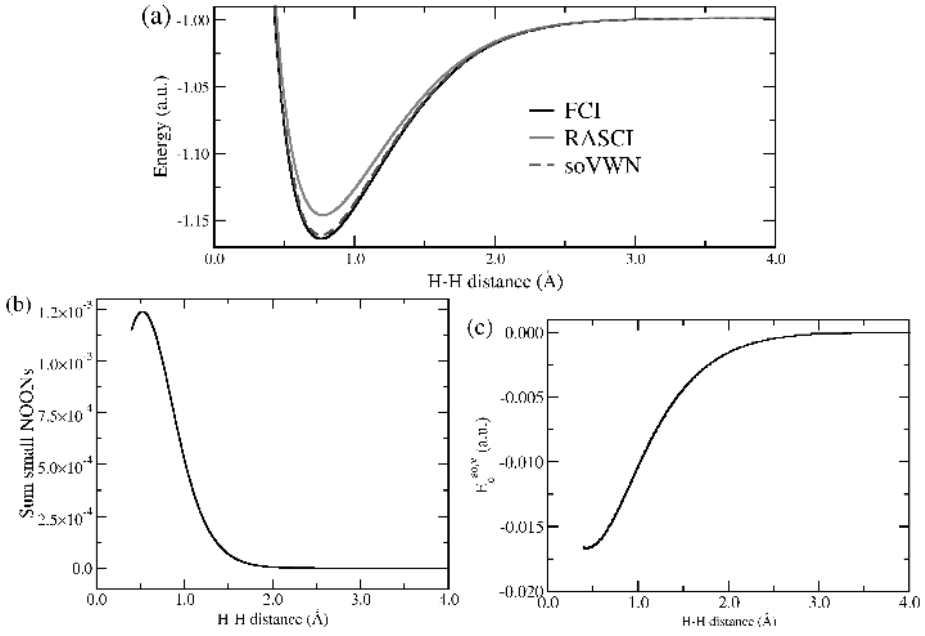


Figure 5.4: (a) Total electronic energy profiles (in a.u.), (b) sum of small NOONs, and (c) RAS-*so*VWN (*so*VWN in short) energy (in a.u.) along the H_2 dissociation.

The improvement in the calculated relative energy between the closed and open electronic structures becomes evident with the computed dissociation energy, as shown in Table 5.1. At

the equilibrium distance, the absence of dynamic correlation in RAS-SF leads to a significant underestimation of the binding energy compared to FCI, notably corrected by RAS-soVWN. It is noteworthy that, in this case, the wave function energy contribution to RAS-soVWN ($\mathcal{E}_0^{\mathrm{WFA},\nu}$) yields nearly identical values to those obtained with RAS-SF, indicating that the soVWN term ($E_c^{so,\nu}$) primarily drives the increase in the computed dissociation energies. In this example, the most favorable outcome is achieved with a threshold of $\nu = 0.01$ and employing Eq. 5.19 for the distribution of the NOONs of the HEG. Larger values, i.e., $\nu = 0.1$, tend to overestimate the correlation energy from NOs with low occupation, while $\nu \leq 0.001$ for RAS-soVWN(S) and smaller for RAS-soVWN(G) lead to vanishing $E_c^{so,\nu}$.

Table 5.1: Dissociation energies (in mHartrees) for H_2 within the cc-pVDZ basis set. These results illustrate three different partitions of the NO space defined through the values for the threshold. $\mathcal{E}_0^{\mathrm{WFA},\nu}$ wave function contribution defined in Eq. 5.11.

method	ν-threshold		
	0.1	0.01	0.001
soVWN(S)	177.3	156.6	148.6
soVWN(G)	192.7	162.2	162.2
$\mathcal{E}_0^{\mathrm{WFA},\nu}$	145.9	148.5	148.5
RAS-SF	141.3		
FCI	163.9		
Exp.[a]	164.5		

[a] Experimental value taken from Ref .[4].

5.6.3 Ethylene torsion

Double bond torsion in ethylene represents the simplest molecular model of a single π-bond breaking. At equilibrium geometry, the two carbon p_z orbitals, perpendicular to the molecular plane, form the bonding π and anti-bonding π^* orbitals. Departure from the planar (D_{2h}) structure through the torsion between the two methylene moieties diminishes the overlap and the interaction between the two p orbitals that become orthogonal at 90° of torsion (zero overlap). At the perpendicular disposition the two frontier orbitals are degenerated and the system exhibits a perfect diradical character. As pointed out in prior works, the accurate description of the torsional potential claims for a balanced treatment of electron correlation effects [389, 390, 391, 392, 393, 394, 395, 396].

Single reference methods struggle to describe the electronic structure for large dihedral angles. At 90° torsion, the wave function should include the $(\pi)^2$ and $(\pi^*)^2$ configurations treated on an equal footing. But the $(\pi^*)^2$ configuration is completely neglected by the restricted Hartree-Fock (RHF) model. As a result, RHF produces a potential curve with an unphysical cusp and a large overestimation of the barrier height (101.5 kcal/mol) compared with the ∼65 kcal/mol estimated from experimental data [397] or by multiconfigurational correlation methods [398] (Table 5.2). This failure of the mean-field solution is carried over to post-HF wave functions such as MP2, CISD or CCSD, or even in (restricted) Kohn-Sham approximations to DFT [391]. Energy profiles are notably improved by spin-polarized (unrestricted) schemes, multireference wave functions, e.g., CASSCF [396], or by the use of SF methods in combination with a high-spin (triplet) reference configuration [394].

In this study, we conduct calculations along the twist mode of ethylene while keeping all other degrees of freedom frozen at the ground state (singlet) geometry, which was optimized at the CCSD(T)/cc-pVDZ level [399]. RASCI calculations were performed using a SF operator

acting on the ROHF triplet state, employing a RAS2 space with 8 electrons distributed over 7 orbitals, and utilizing the cc-pVDZ basis set. The potential energy curves along the torsional coordinate are presented in Figure 11. The RAS-SF wave function yields a smooth energy profile, demonstrating good agreement with prior studies [391, 397, 396]. Notably, the energy gap between the planar and orthogonal conformations (54.5 kcal/mol) avoids the substantial overestimation observed in (restricted) HF and B3LYP calculations. However, it does exhibit a slight underestimation of the torsion barrier by approximately 10 kcal/mol when compared to values obtained with highly correlated methods (VOO-CCD and MRCI) [391, 398]. This underestimation is considerably ameliorated by the soDFT correlation, which introduces differential correlation at the planar structure compared to the barrier maximum (twisted geometry). This improvement appears to be robust with respect to the chosen ν-threshold, particularly in the range of 10^{-3} to 10^{-4} (values below 10^{-4} result in negligible soDFT correction). This suggests a counterbalancing effect between the WFA and soDFT contributions.

Table 5.2: Energy barrier (in kcal/mol) computed for the molecular torsion of ethylene. RAS-soVWN (indicated as soVWN), with values in parenthesis corresponding to the employed sum of occupations threshold.

method	barrier
RAS-SF	54.5
soVWN (0.01)	62.24
soVWN (0.001)	66.71
soVWN (0.0001)	67.23
RHF	109.0
B3LYP	95.3[a]
CASSCF	57.1
VOO-CCD	67.3[b]
MRDCI	62.7[b]

[a]Value reported in Computational Chemistry Comparison and Benchmark Data Base Release 22, May 2022 (https://cccbdb.nist.gov) with the 6-31G(d) basis set. [b]From Ref. [391] with the DZP basis set.

5.6.4 Multiple bond breaking: N_2 dissociation

The dissociation of N_2 involves a multi-step bond-breaking process, transitioning from the ground state electronic configuration at the minimum ($\sigma_g^2\pi_u^4\pi_g^0\sigma_u^0$) to the dissociation limit, ultimately yielding two spin-quartet fragments. The accurate description of the correlation energy in this diatomic system has been revisited several times, in many cases for the purpose of evaluating the performance of correlated quantum chemical methods [400, 401]. Several challenges arise in the theoretical depiction of N_2 dissociation, contingent on the chosen method. These include inconsistencies in the CI expansion [402], accounting for valence-core electron correlations [403], or the effect of basis functions with high angular momenta (e.g., g, h, and i functions) [404].

In the subsequent discussion, we delve into the characteristics of the potential energy surface (PES) concerning N_2 dissociation. This exploration is conducted with the RASCI(h, p) and RAS-soDFT computational approaches, employing a RHF singlet reference and incorporating six electrons and six frontier orbitals ($\sigma, \pi, \sigma^*, \pi^*$) as the RAS2 space, in conjunction with the cc-pVDZ basis set. The computed total electronic energy profiles for the stretching of N_2 are depicted in Figure 5.5a. While the RASCI profile exhibits smooth behavior, the

lack of effective dynamic correlation leads to significantly higher energies compared to the more correlated CIPSI solution. Analysis of the WFA contribution ($\mathcal{E}_0^{\mathrm{WFA},\nu}$) reveals that the utilization of NOs imparts only a marginal effect around the bond equilibrium distance, thereby adding correlation energy relative to RASCI. In contrast, the DFT-corrected RAS-soVWN energies, obtained with a ν-threshold value of 0.001, are energetically lower (by approximately 50 mHartrees), that is, closer to the CIPSI curve, but the energy profile presents several (many) discontinuous steps. This issue emerges solely from the soDFT contribution (Figure 5.5b, red line), while the $\mathcal{E}_0^{\mathrm{WFA},\nu}$ has a smooth profile. Analysis of the small electron occupation entering to soDFT (Figure 5.5b, black line) indicates that the sudden changes in $E_c^{so,\nu}$ are linked to the discontinuous distribution of NOs. Concretely, the steps in the RAS-soVWN energy curve occur at interatomic distances in which the number of NOs included in the soDFT part changes (an addition of one NO when bond distance increases). This change implies a discontinuous "jump" in the ν-value employed for the evaluation of $E_c^{so,\nu}$ (eq 5.21). Fixing the number of NOs included in soDFT to a constant value (instead of the parameter ν) represents a partial solution to this artifact (Figure 21). These results highlight an important drawback of the method in the characterization of PESs, additional profiles computed with larger basis sets can be consulted in Appendix B (Fig. 22).

Dissociation energies (as listed in Table 5.3) are computed by taking the difference between twice the energy of a nitrogen atom (with 3 electrons distributed across 3 orbitals within the RAS2 space) and the energy of N_2 at its equilibrium bond distance. As observed in the cases of H_2 dissociation and ethylene torsion, RASCI notably underestimates the binding energy of N_2 in comparison to highly accurate methods and experimental data [4]. The use of ground state energies computed with NOs having NOONs above the ν-threshold already leads to improved results, yielding dissociation energies similar to those obtained from CASSCF. Furthermore, the inclusion of soDFT correlation significantly enhances the computed binding energy, bringing it closer to values obtained from MRCI and experimental measurements, particularly for ν-thresholds in the range of 0.001 to 0.0001. It is worth noting that larger ν-values tend to overestimate the energy difference between N_2 at equilibrium and twice the atomic energy.

Table 5.3: N_2 dissociation energies (in kcal/mol) computed with CASSCF, RASCI, $\mathcal{E}_0^{\mathrm{WFA},\nu}$, RAS-$so$VWN ($so$VWN in short) with the cc-pVDZ, and compared to reference results from the literature.

method	ν-threshold			
	0.1	0.01	0.001	0.0001
$\mathcal{E}_0^{\mathrm{WFA},\nu}$	193.3	193.3	197.7	202.6
soVWN(S)	250.6	250.6	230.0	202.6
soVWN(G)	276.2	282.6	243.6	202.6
RASCI	185.4			
CASSCF[a]	196.6			
NEVPT2[b]	201.2			
CIPSI	201.4			
MRCI[c]	222.1			
Exp.[d]	225.1			

[a] Active space with 6 electrons in 6 orbitals and the cc-pVDZ basis. [b] From Ref. [152, 168, 153]. [c] From Ref. [400] with the [5s 4p 3d 2f lg] ANO basis set. [d] From Ref. [4].

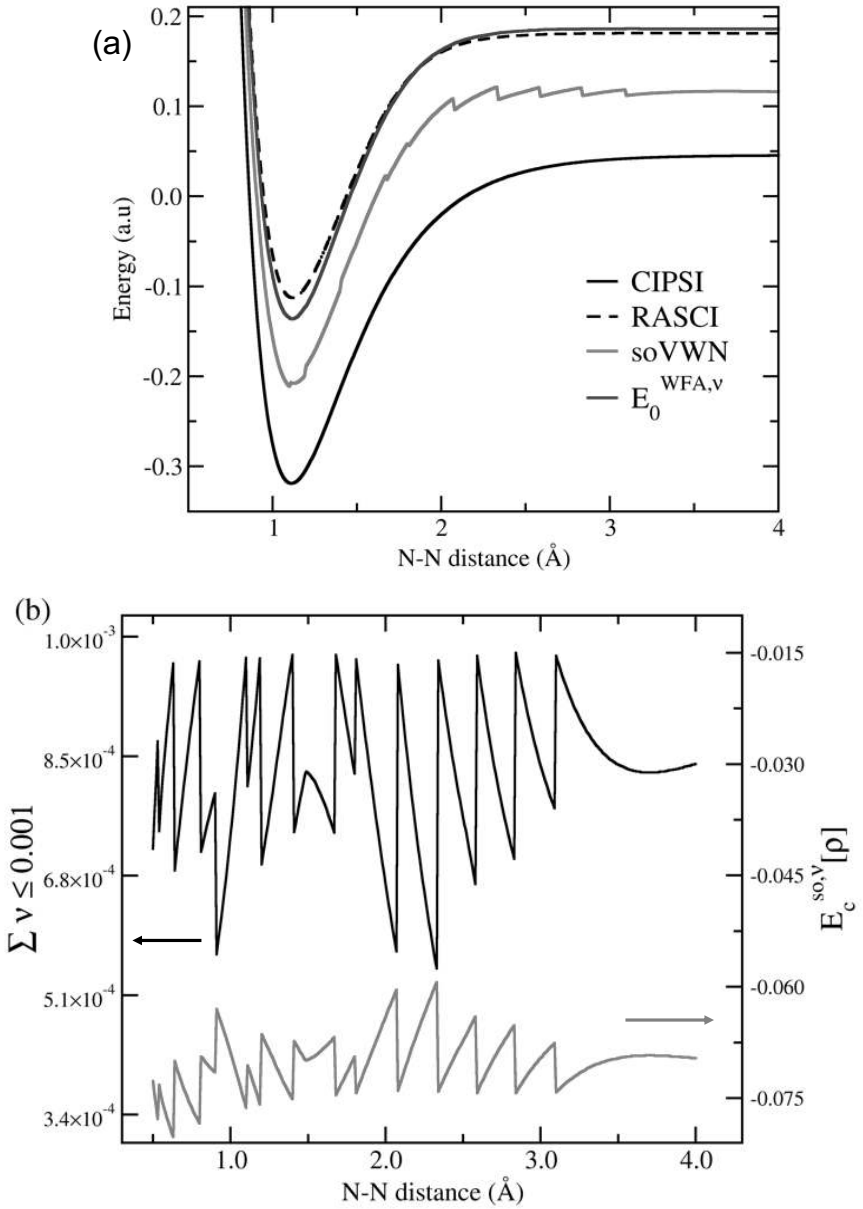


Figure 5.5: (a) Absolute energy profiles along the N_2 dissociation computed at the RASCI (dashed black), RAS-soVWN ($\nu = 0.001$, solid red), $\mathcal{E}_0^{\text{WFA},\nu}$ contribution (solid blue), and CIPSI (solid black) computed with the cc-pVDZ basis set. (b) Sum of NOONs below the threshold $\nu = 0.001$ (black) and $E_c^{so,\nu}$ (red) along the N_2 dissociation.

5.7 Conclusions

In this work have introduced a novel WFT-DFT hybrid approach designed to incorporate DFT correlation into wave function ansatzes by partitioning the orbital space into large and small NOONs, respectively assigned to WFT and DFT correlation. The new method (WFT-soDFT) is based on previous ideas from Savin, but it incorporates different criteria for the splitting of the orbital space and the characterization of the small occupation tail of the HEG distribution. Concretely, we have considered the use of the RASCI wave function approach within the *hole* and *particle* truncation, which provide a simple and rather efficient strategy

to sample the space of small occupation NOs necessary to construct the small-occupation correlation energy functional. Besides the selection of the WFA and correlation functional, the method relies on a separation parameter ν.

The preliminary results here discussed seem to indicate that low ν-values (between $10^{-3} - 10^{-4}$) can provide reasonable correlation energies for the correction of multiconfiguration wave functions of the ground state of medium size molecules. But systematic benchmarking is still necessary to better understand the dependence and universality of the occupation-separation parameter. Moreover, the comparison of WFT-*so*DFT with other WFT-DFT hybrid models, e.g., WFT-*sr*DFT or MC-PDFT, and with multiconfigurational perturbation theory remains to be explored.

These efforts are expected to be carried out in future studies. Since the method requires a good description of the small occupation region of the electron density, it is prone to be used with large basis sets. The results obtained by RAS-*so*VWN in the computation of atomic energies along the He and Be series, the dissociation of single (H_2) and multiple (N_2) bonds, and the torsion of ethylene improve those obtained by the pristine wave function (RASCI).

In some cases, the use of NOs in the WFA already represents an important refinement of the initial wave function, while correlation correction might be important in some situations, e.g., as seen at equilibrium geometries, and insignificant in others, like for the He series with $Z > 2$. The discrete distribution of NOONs obtained with finite basis sets might trigger discontinuities in potential energy surfaces, as seen in the calculation of N_2 dissociation. Despite that this unphysical effect can be partially avoided by fixing the number of NOs in *so*DFT, it represents an important limitation of the method for the characterization of large domains of PESs.

CHAPTER 6

Introducing dynamical correlation into CASSCF wave functions through an effective potential

6.1 Abstract

In this chapter, we explore an alternative method for treating dynamic electron correlation by using an effective correlation potential (V_{eff}) to compensate for the dynamic correlation missing in a CASSCF calculation (CASSCF+V_{eff}). The effective potential is dependent on the electronic separation and is represented as a Gaussian-damped Coulomb interaction. Through the examination of small atomic and diatomic systems, we illustrate the limitations and strengths of this approach. Additionally, based on energy and Coulomb hole analyses for small representative systems, we introduce a variation of the method that involves separating the inter- and intra-atomic contributions to the effective potential. We then explore the impact of different parameters in each contribution. Accurate results suggest that using CASSCF with effective potentials can be a suitable alternative for treating static and dynamic correlation. Improvements in CASSCF energies can be achieved through a judicious selection of parameters in the effective potential. However, several observations are noteworthy. Firstly, qualitative analysis of the Coulomb holes reveals the effects of the effective potential on the electronic distribution. For instance, when the effective potential is repulsive, atomic correlation energies (in general) are worse, but the shape of the respective Coulomb holes can be improved. Secondly, when applied to the dissociation of H_2, CASSCF+V_{eff} Coulomb holes show only minor differences from canonical CASSCF, a behavior attributed to significant contributions from core electrons. Thirdly, extension to other diatomics (N_2, NO^+, O_2, and NH) investigated with different parametrization schemes for the effective potential reveals the distinct effects of the inter- and intra-atomic components of V_{eff}.

6.2 Introduction

Among the plethora of methods available in the field of electronic structure theory, the so-called "correlation factor methods" [25] operate on the premise that short-range correlation, mostly dynamic, cannot be accounted for in a CI expansion (other than FCI), but is addressed

analytically by introducing inter-electronic variables. In fact, the inclusion of a 2-particle potential that directly links the electron correlation into the wave function can lead to solutions of extraordinary accuracy [26, 27, 28, 29], as shown in the calculation of the helium atom by Pekeris [30], the hydrogen molecule by Wolniewicz and Kolos [31] and in explicitly correlated R12 and F12 theories [405, 406, 407]. Early in the literature, two of the most promising correlation factor methods were (1) the method of Boys and Handy and (2) methods based on the correlation hole [25]. The method of Boys and Handy [408, 409, 410, 411] is based on the use of a correlated wave function of the form

$$\Psi = C\Psi_0, \tag{6.1}$$

where Ψ_0 is an approximate, reference wave function (for example, a single Slater determinant), and the factor C is a symmetric function containing the electronic variables,

$$C = \prod_{i>j} F(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j). \tag{6.2}$$

The most basic form of a correlation function involves a dependency solely on the interelectronic distance ($F = F(|\vec{\boldsymbol{r}}_i - \vec{\boldsymbol{r}}_j|)$). However, it is frequently advantageous to incorporate the dependence on position vectors $\vec{\boldsymbol{r}}_1$ and $\vec{\boldsymbol{r}}_2$. In such cases, the extracule and intracule variables, respectively $\vec{\boldsymbol{R}} = \frac{1}{2}(\vec{\boldsymbol{r}}_1 + \vec{\boldsymbol{r}}_2)$ and $\vec{\boldsymbol{s}} = \vec{\boldsymbol{r}}_2 - \vec{\boldsymbol{r}}_1$, are commonly employed. Unfortunately, a correlated wave function in the form of Equation (6.1) leads to a variational energy expression that becomes notably intricate. Realizing this difficulty, Boys and Handy [409, 411] considered a variational procedure based on the selection of an inverse correlation factor (C^{-1}), which vanishes only when $\hat{H}\Psi = \hat{H}C\Psi_0 = EC\Psi_0$, i.e., when the antsatz 6.1 retrieves the exact wave function and energy. When performing this selection, the added correlation energy can be regarded as the expectation value of a (non-Hermitian) "trans-correlated" Hamiltonian [411] of the form

$$\langle\Psi|\,\hat{H}\,|\Psi\rangle = \langle\Psi_0|\,C^{-1}\hat{H}C\,|\Psi_0\rangle\,. \tag{6.3}$$

Selecting appropriate functional forms for $F(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j)$ poses a significant challenge in these methods. A precise choice of F can lead to highly accurate energy estimates, as demonstrated in early studies on representative systems like the Ne atom and the LiH molecule [408, 410]. The complexity involved in finding a suitable form for F under a clear ansatz, leads to further simplifications devoted to approximately account for the dynamical correlation effects. Considering this, in 1987, Roos and co-workers proposed an alternative based on the use of an effective potential (V_{eff}) added to the Hamiltonian [24], the original idea of the authors was to incorporate dynamic correlation into CASSCF calculations. The resulting simplified potential is necessarily empirical, since its parameters are not determined in an *ab-initio* fashion. However, the idea of the method is attractive, since the wave function can be relaxed in the presence of the potential, incorporating dynamical correlation effects within a relatively low computational cost. In the present study, we refer to this approach as CASSCF+V_{eff} and extend the exploration of its performance in describing small atomic and diatomic systems, identifying their general advantages and limitations. Furthermore, we devise and propose a splitting of V_{eff} into *inter*- and *intra*-atomic parts. To that end, we follow Roos and use a CASSCF wave function combined with various proposals for the effective potential. The exploration of CASSCF+V_{eff} performance is achieved through

different parametrization schemes designed to identify the relevance and effects of the V_{eff} constitutive parameters. This chapter is organized as follows: in Sec. 6.3, we introduce further background and theoretical foundation of methods based on the correlation hole, we present Roos' effective potential approximation and expose some of drawbacks and strenghts of the method. Section 6.4 is devoted to show the computational workflow used for CASSCF+V_{eff} calculations, and indicate computational details considered in our calculations. Numerical results are addressed in Sec. 6.5, among them, numerical examples of model atomic systems (Sec. 6.5.1) and H_2 dissociation (Sec. 6.5.2) are used to illustrate the performance of the method, and to justify the development of the proposed V_{eff} splitting described in Sec. 6.6. This proposal is applied to compute potential energy curves (PECs) for other diatomics (sections 6.6.1.1, and 6.6.1.2), where we conduct different parametrization schemes designed to explore the effect of *inter*- and *intra*-atomic terms. Finally, in Sec. 6.7 the main conclusions of the present work are stated.

6.3 Methods Based on the Correlation Hole

When considering a ground state wave function of the form 6.1, we wish to know how the reference pair density, $\rho_2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)$, is modified by the presence of the correlation factor C. A simple approximation to the correlation function $F(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j)$, considered by Colle and Salvetti [412] is

$$F(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j) = 1 - \phi(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j), \tag{6.4}$$

where $\phi(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j)$ is essentially a short-range function, introducing the electron-electron cusp, and can be expressed in terms of the extracule and intracule vectors $\vec{\boldsymbol{R}}$ and $\vec{\boldsymbol{s}}$ (as defined in Section 2.7.5). In deriving ρ_2 from the square of the wave function 6.1, by integrating over $(N-2)$ electronic variables, the form of the general (off-diagonal) elements of ρ_2 reduces to a simple form that depends on the reference 2-RDM

$$\rho_2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2; \vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}'_2) = \rho_2^0(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2; \vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}'_2)\big[1 - \phi(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2) - \phi(\vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}'_2) + \phi(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)\phi(\vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}'_2)\big]. \tag{6.5}$$

The main challenge problem is the selection of the functional form of ϕ able to describe the variation in size and shape of the correlation hole over the electron distribution without violating the requirements for N-representability [413]. Namely:

(I) $\rho_1(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}'_1)$ can be written as the integral of $\rho_2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2; \vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}_2)$ over $\vec{\boldsymbol{r}}_2$ times $1/(N-1)$, i.e. $(N-1)\rho(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}'_1) = \int d\vec{\boldsymbol{r}}_2\, \rho_2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2; \vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}_2)$

(II) as pair correlations do not modify the functional form of $\rho_1^0(\vec{\boldsymbol{r}}_1; \vec{\boldsymbol{r}}'_1)$ computed from the reference wave function (Ψ_0), then we may impose that $\rho_1(\vec{\boldsymbol{r}}_1; \vec{\boldsymbol{r}}'_1) = \rho_1^0(\vec{\boldsymbol{r}}_1; \vec{\boldsymbol{r}}'_1)$.

With these two requirements the following condition is established:

$$\int d\vec{\boldsymbol{r}}_2\, \rho_2^0(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2; \vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}_2)\big[\phi(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)\phi(\vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}_2) - \phi(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2) - \phi(\vec{\boldsymbol{r}}'_1, \vec{\boldsymbol{r}}_2)\big] = 0. \tag{6.6}$$

By means of this, the total energy can be expressed as the reference energy plus a correlation

contribution as

$$E = E_0 + \frac{1}{2}\int d\vec{\boldsymbol{r}}\, d\vec{\boldsymbol{r}}_2 \rho_2^0(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2) \frac{\left[\phi^2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2) - 2\phi(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)\right]}{s} \tag{6.7}$$

which is equivalent to replacing the operator $1/s$ by a screened interaction given as

$$g'(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2) = \frac{\left[1 - 2\phi(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2) + \phi^2(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)\right]}{s}. \tag{6.8}$$

The calculation of correlation energies becomes feasible by a proper form of the function ϕ that must satisfy Eq. 6.6. The form used by Cole and Salvetti is:

$$\phi(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j) = [\Phi(R)(1 + \frac{s}{2a_0})] \exp\left(-\beta^2 s^2\right), \tag{6.9}$$

where a_0 is the Bohr radius, $R = |\vec{\boldsymbol{R}}|$, $s = |\vec{\boldsymbol{s}}|$, and β is a numerical parameter. Within this definition, the function $\phi(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j)$ fulfils all the conditions of the cusp for $\vec{\boldsymbol{r}}_i \to \vec{\boldsymbol{r}}_j$, as demonstrated by Bingel [414]. Note that $\phi(\vec{\boldsymbol{r}}_1, \vec{\boldsymbol{r}}_2)$ has effects on the 2-RDM (Eq. 6.5) modifying the intracule probability (see Eq. 2.180) and the corresponding correlation hole. The function $\Phi(R)$ allows the form of the correlation hole to depend on the mean position of the two particles, and must satisfy the basic cusp condition [196].

The region in space in which $\phi(\vec{\boldsymbol{r}}_i, \vec{\boldsymbol{r}}_j)$ is appreciably different from zero depends on the value of β, this parameter determining the size of the correlation hole. The volume of this region is

$$V = 4\pi \int_0^\infty ds\, \exp\left(-\beta^2 s^2\right) = \frac{\pi^{3/2}}{\beta^3} \tag{6.10}$$

A reasonable value for β can be taken by assuming that this volume is proportional to the volume of exclusion in Wigner's formula for the free electron gas [415],

$$\frac{\pi^{3/2}}{\beta^3} = \frac{k}{\rho}; \quad \beta = \frac{\pi^{1/2}\rho^{1/3}}{k^{1/3}}, \tag{6.11}$$

where k indicates the average number of electrons in V and ρ is the electron density. For a free-electron gas, the parameter β depends only on the electron density, ρ. If we assume that it does so generally, then a dimensional argument suggests $\beta(\rho) = q\rho^{1/3}$ (being q a dimensionless constant).

The relation between $\Phi(\vec{\boldsymbol{R}})$ and β was approximated by Colle and Salvetti as

$$\Phi \sim \frac{\sqrt{\pi}\beta}{1 + \sqrt{\pi}\beta} \tag{6.12}$$

As β varies from 0 to ∞, Φ ranges from 0 to 1. In fact, the ρ dependence in $\beta(\rho)$ is a significant drawback, introducing computational complexities for practical uses of the correlation function ϕ (Eq. 6.9). However, the general validity of this approach has been confirmed in early studies [412, 25] through application to the free-electron gas model, where it is emphasized that such a function provides a good representation of the spherically averaged correlation hole.

6.3.1 The Roos CASSCF Effective Potential Method

Following correlation factor methods, Roos noted that by using the ansatz from Boys (Eq. 6.1), the time independent Schrödinger equation can be written in the form

$$(\hat{H} + C^{-1}[\hat{H}, C])\Psi_0 = E\Psi_0, \tag{6.13}$$

from which the second term in the left hand side can be taken as an effective potential,

$$V_{eff} = C^{-1}[\hat{H}, C]. \tag{6.14}$$

This observation introduces de idea of treating DC effects by modifiying the Coulomb repulsion term in the Hamiltonian, provided that Ψ_0 is a MCSCF wave function capable of a qualitatively correct description of NDC. As noted before, the main problem here is to give an appropriate expression for this potential, which translates into a proper form for the correlation factor F from Eq. 6.4. In fact the idea of Roos can be considered as a further simplification of the formulation of Cole and Salvetti. By using the Colle and Salvetti's function ϕ (Eq. 6.9, which is written in terms of function $\Phi(\vec{\boldsymbol{R}})$) into Eq. 6.4 and using the corresponding form of C (Eq. 6.2), the effective potential becomes too complicated for practical uses. However, as argued by Roos, the function $\Phi(\vec{\boldsymbol{R}})$ is assumed to depend only weakly on $\vec{\boldsymbol{R}}$, then for small inter-electronic distances the leading term of the effective potential can be written in a simple form

$$V_{eff} = -\mu \sum_{i<j} \frac{e^{-\nu r_{ij}^2}}{r_{ij}}. \tag{6.15}$$

A justification for this simplification lies in the fact that the energy of the approximate wave function Ψ_0 (a CASSCF wave function) already contains a significant portion of NDC. Therefore, only a small fraction needs to be added as a correction through the interelectronic dependence of the effective potential. The total energy becomes

$$E = \langle \Psi_0 | \hat{H} - \mu \sum_{i<j} \frac{e^{-\nu r_{ij}^2}}{r_{ij}} | \Psi_0 \rangle \,. \tag{6.16}$$

The energy correction introduced by the effective potential can be considered as the expectation value of a Coulomb operator damped by a Gaussian term. The parameters μ and ν resemble the multiplicative factor and the parameter β from the more elaborate form given in Eq. 6.9. In Roos' formulation, these parameters were chosen to inherit a similar interpretation as that given by Cole and Salvetti, representing the depth (μ) and the radius (ν^{-2}) of the Coulomb hole generated by V_{eff}. The most appealing virtue of this selection is that the effective potential allows the approximate wave function to relax under the influence of the (simplified) correlation term.

In practice, the implementation of the effective potential involves the computation of corresponding electronic repulsion integrals (ERIs). In Sec. 6.6, we show the form of the

fundamental ERI, and provide a brief description about it. The general solution for the required ERIs in the CASSCF+V_{eff} method has been derived and can be consulted in Appendix C.

6.3.2 Drawbacks and Strengths of the Method

The major problem with Roos' proposal is the selection of the parameters. Since formally they must depend on the local density $\rho(\vec{r})$, the simplicity of the potential is not retained anymore. Alternatively, Roos proposed a parametrization based on the use of the average electron density,

$$\overline{\rho} = \frac{\int \rho^2(r)dr}{\int \rho(r)dr}, \tag{6.17}$$

based on the correlation energy results of a series of f [24]. Fitting the results to the empirical values reported by Clementi [416, 417], Roos established a linear relation between μ and ν as

$$\mu = \mu_0 + p\nu, \tag{6.18}$$

where μ_0 and p are used to establish the functional dependence of μ and ν on the average electron density assumed as

$$\mu = x_1\overline{\rho}^{x_2}, \tag{6.19}$$

$$\nu = (x_3 + \Delta x_4)\overline{\rho}^{x_5}, \tag{6.20}$$

where $x_1 - x_5$ are constant numbers determined through a minimization of the least square errors of Eq. 6.18. The parameter Δ appearing in Eq. 6.20 is introduced to account for the amount of correlation already introduced into the CASSCF wave function. In principle, the radius of the Coulomb hole should go to zero when Ψ_0 approaches the exact wave function. The determination of Δ involves an implicit calculation and integration of the two-body density matrix (ρ_2).

Thus, the main drawbacks of Roos' proposal can be summarized in: the laborious parametrization scheme, the fact that the secondary parameters, μ_0 and p, vary for each electronic state of an atom or ion, and the additional cost involved in the integration of ρ_2. The parameters' dependence on the average electron density introduces additional difficulties.

However, the effective potential energy correction is relatively simple and does not require large amounts of computational resources. Despite its publication in 1987 (thirty-six years ago), this proposal has not received much attention from the theoretical chemistry community. Nonetheless, three key points are worth highlighting:

- (1) It was demonstrated that the effective potential can be added to the Hamiltonian to retrieve the exact enegy (theoretically). Together with a parametrization based

on experimental correlation energies, and the use of a large enough basis set, the potential retrieves highly accurate energies from approximate wave functions for atoms.

- (2) There is no clear hint of how to construct the proposed potential. However, it can be seen as an approximation to the expectation value of a (non-Hermitian) 'trans-correlated' Hamiltonian [411, 25]. The electron-electron interaction is given by $(1 - \mu \exp\{-\nu r_{12}^2\})/r_{12}$, where r_{12} is the interelectronic distance, ν and μ being adjustable parameters with the implicit dependence on the average density and 2-RDM.
- (3) Based on dimensional analysis and through parametrizations, a relationship between μ and ν could be established.

In fact, Roos did not provide an exhaustive study of the behavior of electron correlation in the presence of the effective potential for systems beyond first-row atoms. Additionally, alternatives for the parametrization scheme or the dependency of parameters on quantities other than the average electron density were not provided. Computational power available at his time was a significant limitation; however, nowadays, accurate electronic structure calculations can be more easily performed, especially for small- and medium-sized systems.

6.4 Methods

6.4.1 Computational Workflow

As previously mentioned, the use of the effective potential (Eq. 6.15) and evaluation of the total energy (Eq. 6.16), involves evaluation of the attenuated Gaussian damped ERIs. In this work, we use an *in-house* code designed to perform this task. The evaluation algorithm, is not based on schemes using recurrence relations [104] or numerical integration quadrature [418]. Instead, it is based solely in the direct computation of each ERI provided the general form in Appendix C. The common bottleneck here is the solution of the Boys function ($F_n(x)$) involved in the ERI evaluation,

$$F_n(x) = \int_0^1 dt\, t^{2n} \exp\left(-xt^2\right). \tag{6.21}$$

Since there is no analytical solution to solve this function (except for very specific cases) [102, 103], our algorithm numerically evaluates it through the use of an efficient open-source library (libboys) based on the use of precalculated values and a recursive formula up to six orders [419].

Once ERIs have been obtained, CASSCF+V_{eff} energy calculation can be done in a simple way by a suitable modification of the electronic Hamiltonian. A practical and general procedure for this can be schematically described as follows:

1. Initial HF calculation to obtain reference state for CASSCF calculation and mean-field ERIs.
2. Separate calculation of modified-ERIs within the same basis set.
3. Compute the total ERIs as the sum of the original ERIs plus Gaussian damped ERIs (multiplied by the corresponding μ parameter).

4. Perform the CASSCF calculation with the total ERIs.

5. Obtain the total energy by standard CASSCF minimization procedures.

The current implementation of this scheme consists in a python package that combines the *in-house* integral solver with PySCF code [420]. As such, it enables flexible calculations within an easy-to-manage interface.

6.4.2 Computational Details

Throughout this work we employ different parametrization schemes to explore and elucidate the relevance and effects of the different parameters of the potential. It is worthwile to highlight that μ parameters in V_{eff} are those who will define a repulsive (positive sign) or attractive (negative sign) potential. As will be indicated in subsequent subsections, we will opt for the exploration of both types of potential. Otherwise specified, all calculations reported in this work were performed within the Dunning's correlation-consistent cc-pVDZ basis set [421].

Also, we make use of different analysis tools to elucidate different properties of the systems under study, such as, Coulomb holes, delocalization indices (here indicating the number of electrons that are shared or exchanged between two atoms) [422, 423], atomic volumes and correlation indices [252]. The purpose of their use is to identify the possible effects of V_{eff} over pristine CASSCF results, to establish which of the parameters in V_{eff} can be related to a given effect, and to identify the behavior of DC and NDC along diatomic PECs.

Coulomb holes were computed under Coulson's definition [263, 264, 265, 259]. The density matrices and the intracule probability distributions have been obtained with the in-house DMN[424] and RHO2_OPS codes [425], respectively, the latter using the algorithm of Cioslowski and Liu [426]. For ground state dimers, delocalization indices and atomic volumes were computed from Quantum Theory of Atoms in Molecules (QTAIM) [427, 428] calculations implemented in AIMAll software [429] starting from molecular wave function data computed with Gaussian 16 [430].

In this work, we explore the capabilities and performance of the CASSCF+V_{eff} method in computing the correlation energies of first-row atoms in their ground state (B, C, N, O, F, Ne), and in the calculation of the PECs of H_2, N_2, and NO^+ singlet ground state, and triplet ground state O_2, NH molecules.

6.5 Numerical Results

6.5.1 Performance in atomic systems

When applied to atomic systems, Roos and co-workers demonstrated that the effective potential strategy can improve the dynamical correlation energy, showing that some relationship between μ and ν parameters could be established, so that the number of parameters could be effectively reduced. Additionally, Roos' results showed that, within a proper parametrization, highly accurate correlation energies can be achieved. However, they gave no clear hint about how to obtain the correct parametrization.

Here, with the aim to show the performance of the potential over a more extended set of values for the parameters in the effective potential, we extend Roos' study on atoms. Similar to Roos' study, full-valence CASSCF+V_{eff} calculations were carried by varying ν-values in

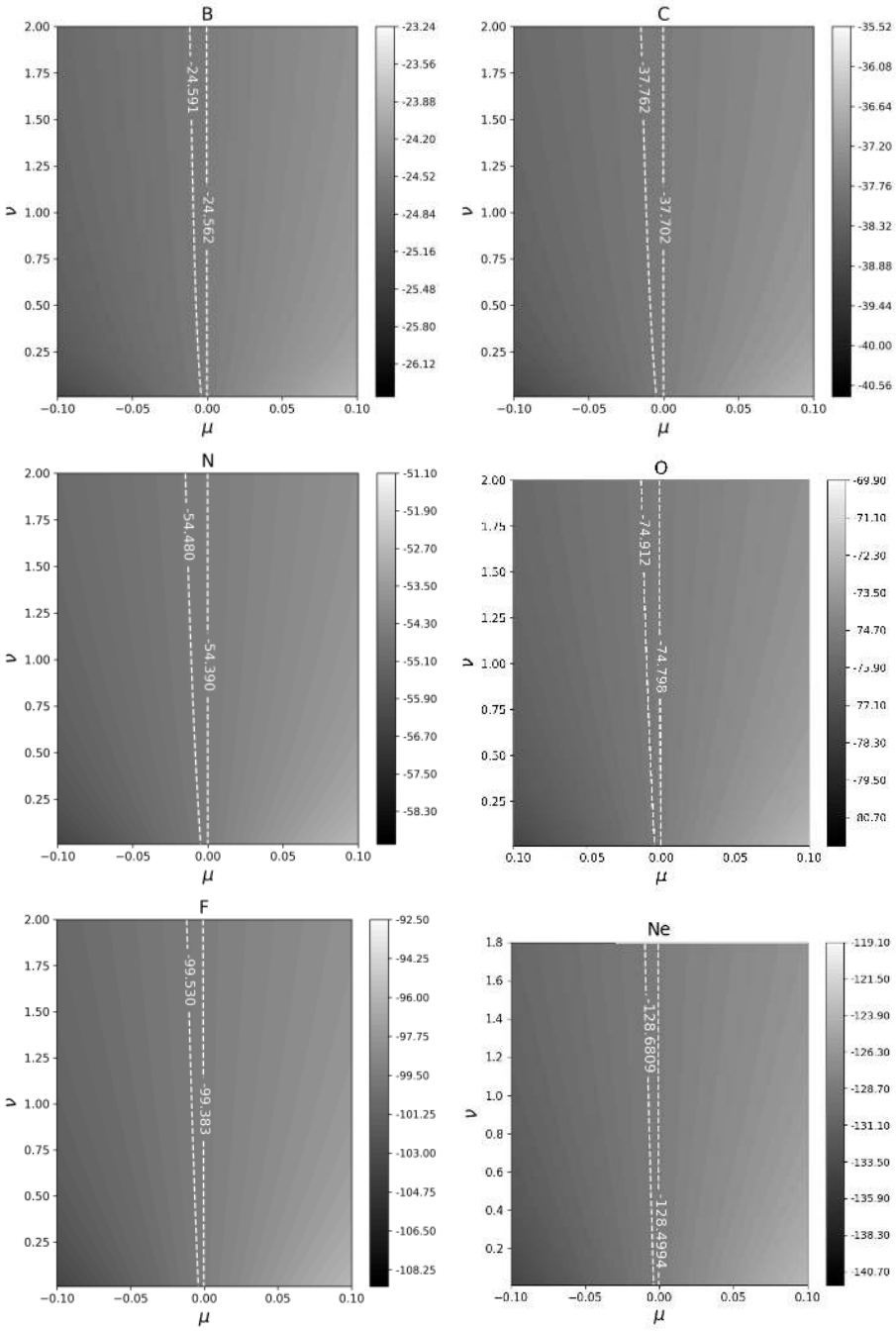


Figure 6.1: Atomic total energies (in a.u.) computed within the CASSCF+V_{eff} represented as contour plots resulting from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot. The CASSCF energy lies in a vertical line at $\mu = 0$.

the range from 0.01 to 2.00 in steps of 0.01, and μ-values in the range from -0.200 to 0.200 in steps of 0.001 (giving a total of 80 000 individual calculations per atom). The total energies computed in this way (restricted to μ-values $\in [-0.10, 0.10]$), are depicted as contour plots in Figure 6.1, where dashed lines are used to indicate the (μ, ν) pairs retrieving FCI (lower values) and CASSCF (higher values) results.

Similar plots for the corresponding total and correlation energies of ions can be consulted

in the Appendix C. Computed total energies exhibit a consistent trend across all atomic systems. Significant energy variations occur at fixed ν values, while changes in ν when keeping μ constant show minimal variation, indicating the relative independence of these two parameters. The behavior at $\mu \to 0$ seems to be in agreement with the linear relation between the parameters considered by Roos (Eq. 6.18). The correlation energies computed as $E - E_{HF}$ follow a similar trend (see Appendix C, Fig.25).

Our findings suggest that FCI/cc-pVDZ energies can be achieved by the CASSCF+V_{eff}. This level of accuracy can be obtained at small μ values, in the range of -0.02 to 0.02, but energies tend to be substantially overestimated (or underestimated) quickly at larger negative (positive) values of this parameter. This trend also holds in the case of ions (see Appendix C, Figs. 27, 28 and 29). The results highlight the possibility of computing highly accurate energies for atoms by carefully selecting the parameters in the effective potential. To achieve this, we use a linear relation of the form $\nu = a\mu + \mu_0$ and performed a data fitting from those combinations of (μ, ν) yielding correlation energies within approximately 1 kcal/mol above the FCI result at most. ν values and corresponding correlation energies computed in this way (taking at lowest value of $\mu = -0.006$) can be consulted in Table 6.1.

Table 6.1: Full valence CASSCF, FCI, and CASSCF+V_{eff} ($\mu = -0.006$) correlation energies for first row atoms and ions (in a.u.). Calculations performed with the cc-pVDZ basis set. Corresponding ν values computed from the data fitting ($r^2 > 0.99$).

Correlation Energy				
Atom	CASSCF	FCI	CASSCF+V_{eff}	ν
B(^{2}P)	-0.0338	-0.0640	-0.0632	0.2623
C(^{3}P)	-0.0187	-0.0795	-0.0786	0.1241
N(^{4}S)	0.0000	-0.0917	-0.0909	0.1632
O(^{3}P)	0.0000	-0.1242	-0.1233	0.2567
F(^{2}P)	0.0000	-0.1577	-0.1572	0.4157
Ne(^{1}S)	0.0000	-0.1921	-0.1886	0.6848
C^+(^{2}P)	-0.0415	-0.0726	-0.0708	0.9502
N^+(^{3}P)	-0.0220	-0.0833	-0.0811	0.4540
O^+(^{4}S)	0.0000	-0.0911	-0.0891	0.4853
F^+(^{3}P)	0.0000	-0.1203	-0.1176	0.6803
Ne^+(^{2}P)	0.0000	-0.1507	-0.1479	0.9678
C^-(^{4}S)	0.0000	-0.0867	-0.0861	0.0340
O^-(^{2}P)	0.0000	-0.1573	-0.1560	0.1731
F^-(^{1}S)	0.0000	-0.1948	-0.1934	0.3131

To show the performance of this parametrization, computed correlation energies within CASSCF+V_{eff} and their comparison with FCI and pristine CASSCF are compared in Table 6.1. As shown, calculations with the parametrized effective potential follow a close agreement with FCI. In a similar fashion, the corresponding energies using parametrized potentials for cations and anions, were used to calculate the ionization potentials (IPs) and electron affinities (EAs) reported in Table 6.2. While computed IPs are close to experimental values, obtained EAs are close to the FCI value but largely deviate from the experimental value in all cases, feature that can be attributed to the quality and size of the basis set. These overall results demonstrate the capability of the method to deliver accurate results from an optimal parametrization.

While obtained energies can be tuned through the effective potential to deliver an accurate result, it can be argued that during the variational procedure, the addition of the effective potential may affect the electronic distribution of the system. In general it can be expected that, as the strength of the potential (measured by the magnitude of μ) becomes significant, larger effects on the electronic distribution may arise. In order to estimate these effects, Coulomb holes for C, F^- and Ne atomic singlet states (^1S) were computed from both, the CASSCF and CASSCF+V_{eff} wave functions. In these calculations, positive-μ (repulsive) and negative-μ (attractive) effective potentials were considered, and ν values from previous parametrization were used.

Table 6.2: Atomic ground state ionization potentials and electron affinities in eV.

Atom	Method	IP	EA
C	CASSCF	10.1763	-0.6093
	CAS+V_{eff}	11.0097	0.1143
	FCI	10.9851	0.0955
	Exp.	11.25	1.12
N	CASSCF	13.3661	-3.3399
	CAS+V_{eff}	14.2288	-2.6504
	FCI	14.1927	-2.5472
	Exp.	14.54[a]	
O	CASSCF	11.9528	-1.8849
	CAS+V_{eff}	12.8819	-0.9823
	FCI	12.8550	-0.9850
	Exp.	13.61[a]	1.4611[c]
F	CASSCF	15.6964	-0.1600
	CAS+V_{eff}	16.7758	0.8354
	FCI	16.7127	0.8505
	Exp.	17.42 [a]	3.45 [b]
Ne	CASSCF	19.7688	
	CAS+V_{eff}	20.8766	
	FCI	20.8962	
	Exp.	21.56 [a]	

[a] From Ref. [431] [b] From Ref. [432] [c] From Ref. [433].

Coulomb holes for atomic singlets (Ne and F^-) are depicted in Figure 6.2. A comparison of these results reveals that the sign of the effective potential directly influences the Coulomb holes. Here, it is worth noting that these Coulomb holes do not originate from the exact wave function ($C\Psi_0$) but rather belong to the corresponding Ψ_0 optimized in the presence of V_{eff}. In this sense, Coulomb holes account for the overall effect of V_{eff} on the electronic distribution, which is linked to the corresponding change in energy.

An attractive ($\mu < 0$) effective potential effectively reduces the total energy, as anticipated with the inclusion of dynamic correlation. However, in this scenario, the average distance between electron pairs is reduced, contrary to the expected effect of electron correlation

inclusion. Therefore, while the energy improves, the wave function deteriorates. Conversely, a repulsive effective potential (Coulomb hole in dashed lines) exhibits an improved behavior, consistent with FCI or the pristine CASSCF. With this choice, energies for atoms are clearly worsened, but Coulomb holes are improved, especially at short interelectronic distances. Additional results for the C atom can be found in the Appendix C.

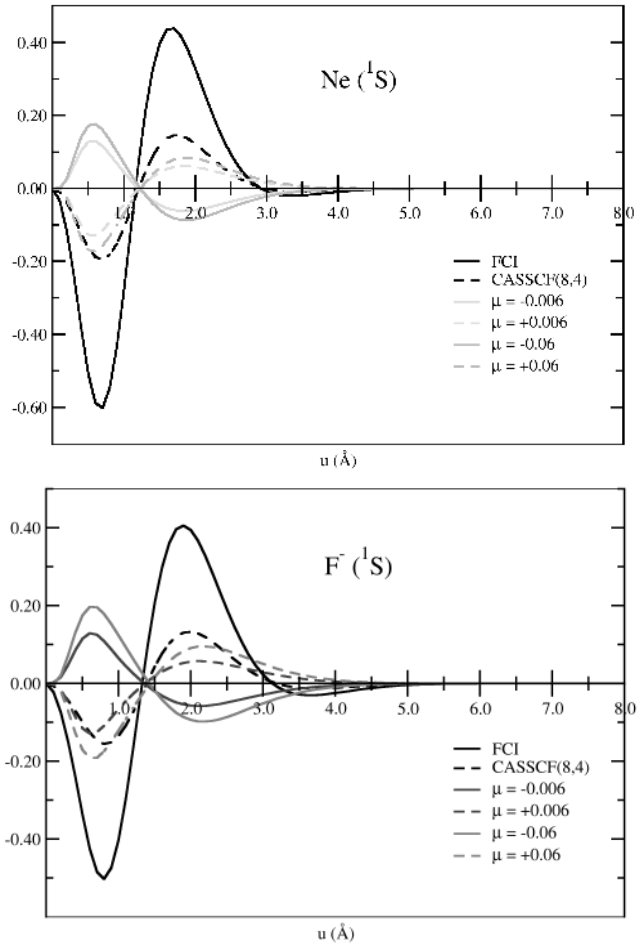


Figure 6.2: Coulomb holes computed for singlet atomic systems at FCI/cc-pVDZ (continuous black line), CASSCF (dashed black line), and CASSCF+V_{eff} with different μ (corresponding ν taken from parametrization with FCI).

6.5.2 Describing H_2 dissociation

In extending the application of CASSCF+V_{eff} to diatomics, we start with the exploration of method's performance in the description of the H_2 PEC. As exposed in previous chapters (see Sec. 4.3.3 and Sec.5.6.2), this is a system for which the ground state experiences continuous changes along the H-H stretching; dynamic and static correlations play a significant role at different regions of the PEC. For instance, at equilibrium, the ground state corresponds to a closed-shell singlet, and dynamic correlation governs over static part. As the separation of the atoms increases, the ground state experiences the conversion to an open-shell singlet with two effectively unpaired electrons.

Near equilibrium, a single Slater determinant is able to give a good description of the singlet state. In consequence, approaches such as HF or KS-DFT provide a sufficiently accurate

description of the system. As the molecule is stretched, the σ and σ^* frontier orbitals become nearly degenerated and partially occupied. The wave function is no longer well described by a single Slater determinant, and nondynamic correlation becomes dominant at dissociation.

The correct description of PECs involves a balanced treatment of DC and NDC contributions. In the case of H_2, DC must increase at bonding region and decrease progresively as the H-H distance is increased. In Figure 6.3, we depict the H_2 PEC computed with different levels of theory. It is particularly satisfactory the result obtained within the CASSCF+V_{eff}, which is in good agreement with FCI dissociation energy, representing an improvement over the CASSCF(2,2) solution.

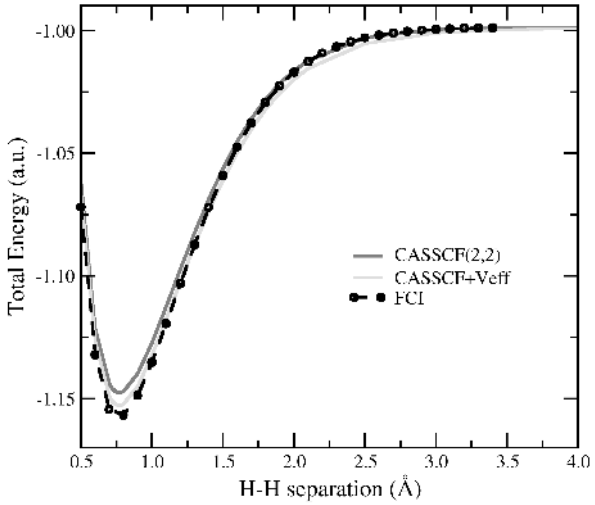


Figure 6.3: H_2 PEC comparison of different methods. This CASSCF+V_{eff} accurate result was obtained within the combination of parameters: $\mu = -0.02$ and $\nu = 0.50$, obtained by computing dissociation energies with different (μ, ν) combinations, and selecting those closest to the FCI result.

The computed CASSCF+V_{eff} PEC in Fig. 6.3 is an illustrative result chosen from a set of calculations performed by taking in consideration combinations of (μ, ν) parameters, such that μ assumes values in the range $[-0.20, 0.20]$ in steps of 0.001, and exponential ν parameter is chosen among these values: $0.05, 0.10, 0.50, 1.00, 1.50, 2.00$, all the points in the PEC were computed using the same (μ, ν) combination. In all cases, it is observed that independently of ν, values near $\mu = 0.020$ value yield the lowest error in the dissociation energies with respect to FCI/cc-pVDZ (see Table 6.3).

A comparison of CASSCF+V_{eff} Coulomb holes computed with a repulsive and attractive V_{eff} ($\mu = \pm 0.020$, $\nu = 0.50$) is depicted in Fig. 6.4. When we apply the Roos' effective potential method to H_2 at equilibrium geometry, the Coulomb holes do not deteriorate as in the case of atoms. In this case neither, the change of sign in μ nor variations of ν seem to have a large effect on the Coulomb hole with respect to CASSCF (related results can be consulted in Appendix C, Fig. 34). The latter due to the lack of core electrons in H_2. Consequently, it serves as a scenario where the dynamic correlation introduced by the effective potential originates solely from *inter*-atomic electron pairs. Therefore, we can expect the observed changes in Coulomb holes of atoms to be similar to those we will observe for the inner electrons of diatomics.. Additional results obtained with different (μ, ν) combinations exhibit a similar behavior, as detailed in Table 6.3. The corresponding Potential Energy Curves (PECs) are depicted in Figure 6.5, where we compare cases in which the (μ, ν) parameters in

the effective potential yield accurate energies close to the FCI level. Notice also that μ for a potential of electron pairs between atoms ($\mu = 0.2$) take very different value than we had found for inner electrons ($\mu = -0.006$, see Table 6.1).

Table 6.3: Most accurate CASSCF+V_{eff} ($\mu = -0.02$) dissociation energies and percentage of error with respect to FCI/cc-pVDZ dissociation energy for H_2 system.

ν	E_{diss} (eV)	% Error
0.05	4.316	-0.656
0.10	4.280	-1.383
0.50	4.179	-3.718
1.00	4.130	-4.830
1.50	4.109	-5.330
2.00	4.096	-5.616
CASSCF(2,2)	4.051	-6.664
FCI	4.340	

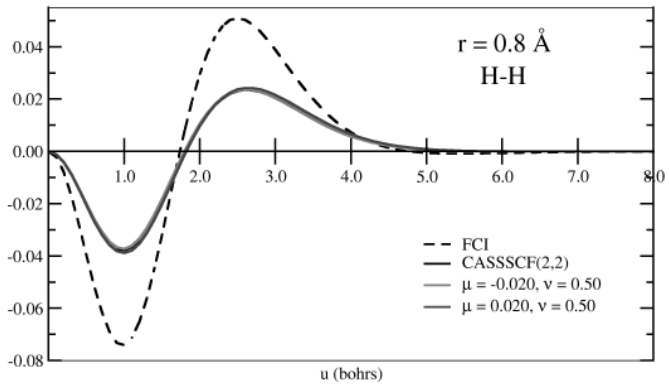


Figure 6.4: H_2 Coulomb holes at equilibrium geometry, computed within CASSCF(2,2), FCI and CASSCF(2,2)+V_{eff} with the cc-pVDZ basis set. The values for the parameters in the effective potential are depicted in the figure.

Furthermore, the results from Table 6.3 suggest that as ν decreases (from 2.00 to 0.05), the dissociation energies exhibit improvement compared to the CASSCF result.

The attenuation of V_{eff} is modulated by ν; the larger this parameter, the more rapidly the potential approaches zero with interelectronic separation r_{ij}. Thus, a large ν value allows for a less extensive effect of the potential over space, causing the CASSCF+V_{eff} energy to decay faster towards the CASSCF result as the ν value grows. On the other hand, a more extended insertion of dynamic correlation over space is achieved with smaller ν values, improving the energies over more points in the PEC.

However, as illustrated in Figure 6.5 (a), in the intermediate region (between 2.00 and 3.00 Å), the PEC is significantly influenced by the action of the effective potential ($\nu = 0.10$). This influence causes the energy to lie lower than the FCI solution, and this effect intensifies as ν decreases. Therefore, in contrast to atomic systems showing a light dependence on ν, we can anticipate that for processes involving the separation of electron pairs, this parameter must have a more pronounced effect on the energy. The latter suggests that inner electron pairs and inter atomic pairs should be treated differently within this approach.

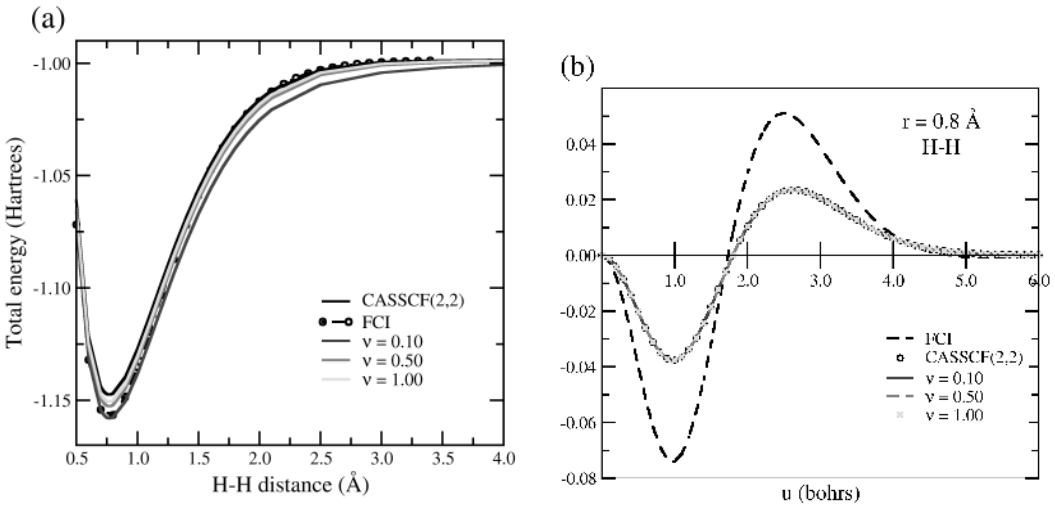


Figure 6.5: (a) Most accurate H_2 PECs estimated from parametrization based on the dissociation energies. Effect of different ν values. (b) Corresponding Coulomb holes at equilibrium geometry ($r_{H-H} = 0.8$ Å). In all cases, $\mu = -0.02$ and the cc-pVDZ basis set were used.

6.6 Splitting the Roos' Effective Potential

From findings exposed previously, it is noted that, for atomic systems, the primary contribution to V_{eff} is $intra$-atomic, while for H_2, it is exclusively $inter$-atomic. These observations have motivated the decision to split the effective potential into these two contributions, as indicated in Eq. 6.22 (see below). In the subsequent sections, we direct our efforts towards describing the PECs for diatomic molecules containing atoms heavier than hydrogen. Specifically, we choose to model the dissociation of representative diatomic molecules, including the isoelectronic singlet ground state N_2 and NO^+, as well as the triplet ground state O_2 and NH diatomics.

Since for chemical analysis and photophysics we are concerned about relative energies, not absolute energies, it is more important to describe correctly the *inter*-atomic interactions than the *intra*-atomic ones for energy analysis; while for the properties it is important not to deteriorate the quality of the CASSCF wave function and improve it, whenever possible. With the aim to split the dynamic correlation into these two terms, and to obtain balanced results we propose to split Roos' potential into

$$\hat{V}_{eff} = \mu_{intra} \sum_{i<j} \frac{e^{-\nu r_{ij}^2}}{r_{ij}} + \mu_{inter} \sum_{p<q} \frac{e^{-\nu' r_{pq}^2}}{r_{pq}} = \hat{V}_{eff}^{i} + \hat{V}_{eff}^{o}, \tag{6.22}$$

where i and o refer to the intrapair electrons and interpair electrons, respectively. The first term represents an overall effect on the *intra*-atomic contributions weighted by the parameters (μ_{intra}, ν) and defined in such a way that includes mostly pairs of core electrons. The second term belongs to the *inter*-atomic contributions, weighted by the corresponding pair parameters (μ_{inter}, ν'). This term includes the rest of pairs of electrons not accounted for in the *intra*-atomic part. The expectation value of this effective potential (Eq. 6.22) will apply using the same wave function with both μ_{intra} and μ_{inter}. In practice, since electrons are indistinguishable, the application of the effective potential discriminates the respective inter-

and intra-pairs, identified by the corresponding two-electron Gaussian damped integrals, arising from

$$\left\langle \phi_A\phi_B \left| \frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}} \right| \phi_C\phi_D \right\rangle . \tag{6.23}$$

Atomic orbitals, ϕ_A, ϕ_B, ϕ_C and ϕ_D are contracted Gaussian functions centered on nuclear centers, associated to $\vec{\boldsymbol{A}}$, $\vec{\boldsymbol{B}}$, $\vec{\boldsymbol{C}}$ and $\vec{\boldsymbol{D}}$ vectors, respectively. In particular the fundamental ERI of this type corresponds to the case in which the atomic orbitals are of zero angular momentum, in this case the integral reduces to

$$\begin{aligned} \left\langle ss \left| \frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}} \right| ss \right\rangle &= \left(\frac{2\pi^{5/2}}{(\nu\gamma_P + \nu\gamma_Q + \gamma_P\gamma_Q)(\gamma_P + \gamma_Q)^{1/2}} \right) \\ \times \exp\left(-\frac{\gamma_P\gamma_Q\nu\overline{PQ}^2}{\nu\gamma_P + \nu\gamma_Q + \gamma_P\gamma_Q} \right) &\cdot F_0\left(\frac{\overline{PQ}^2\gamma_P^2\gamma_Q^2}{(\nu\gamma_P + \nu\gamma_Q + \gamma_P\gamma_Q)(\gamma_P + \gamma_Q)} \right), \end{aligned} \tag{6.24}$$

where $F_n(X)$ is the so-called Boys function (defined for the general case in Appendix C), $\overline{PQ}$ is the distance between the reduced centers, $\vec{\boldsymbol{P}}$ and $\vec{\boldsymbol{Q}}$, written in terms of the exponential coefficients associated to each contracted function, respectively, α_A, α_B, α_C, α_D as

$$\begin{aligned} \vec{\boldsymbol{P}} &= \frac{\alpha_A\vec{\boldsymbol{A}} + \alpha_B\vec{\boldsymbol{B}}}{\gamma_P}, \\ \vec{\boldsymbol{Q}} &= \frac{\alpha_C\vec{\boldsymbol{C}} + \alpha_D\vec{\boldsymbol{D}}}{\gamma_Q}, \end{aligned} \tag{6.25}$$

and $\gamma_P = \alpha_A + \alpha_B$, $\gamma_Q = \alpha_C + \alpha_D$. The splitting of the effective potential showed in Eq. 6.22, is based on the selection of those ERIs from primitives sharing the same center ($A = B = C = D$), this conforms the *intra*-atomic term, while the rest contains those contributions with at least one primitive on a different center, giving the *inter*-atomic term. For diatomic molecules, this is equivalent to the two following conditions:

$$\left\langle \phi_A\phi_B \left| \frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}} \right| \phi_C\phi_D \right\rangle_{intra} = 0, \quad \text{unless } A = B = C = D \tag{6.26}$$

$$\left\langle \phi_A\phi_B \left| \frac{\exp\left(-\nu' r_{ij}^2\right)}{r_{ij}} \right| \phi_C\phi_D \right\rangle_{inter} = 0, \quad \text{if } A = B = C = D \tag{6.27}$$

For larger molecules, three atom terms and higher would be neglected. Notice that, compared to Roos' original potential, Eq. 6.22 sensibly reduces the number of additional two-electon integrals needed when restricting the calculation to one or the other component.

6.6.1 Studies in diatomics beyond H_2

To optimize CASSCF+V_{eff} results, we explore two parametrization schemes designed to approach accurate energies with optimal μ_{intra}, μ_{inter}, and ν parameters in V_{eff}. The first parametrization scheme (referred to as PS1) involves the calculation and comparison of dissociation energies with respect to reference FCI values [2]. Similar to the selection of accurate PECs described before for H_2, in this case, we consider optimal (μ_{inter}, ν) and (μ_{intra}, ν) pairs, those providing D_e values with percent errors $\leq$10% relative to FCI dissociation energies. To achieve this, we compute the CASSCF+V_{eff}/cc-pVDZ total energies computed at large atomic separations (20 Å) and compute the corresponding dissociation energies by subtracting the energy at CASSCF equilibrium geometries. To establish an acceptable range of parameter values, we vary μ_{intra} and μ_{inter} in the range of $[-0.03, 0.03]$, with steps of 0.001. The ν-exponential parameter is held fixed for both intra- and inter-contributions.

PS1 allow us to obtain optimal parameters such that CASSCF+V_{eff} energies become of FCI quality, however this might not be sufficient for a good quality PEC. Therefore, motivated by pursuing for complete PECs of FCI quality, we explore the second parametrization scheme (referred to as PS2), based on the elucidation of optimal parameters for each point in the PEC with respect to FCI/cc-pVDZ energies computed within the perturbatively selected configuration interaction scheme (CIPSI) implemented in the Quantum Package [113]. For each point in the PECs, CASSCF+V_{eff} energies are computed using the same range of μ_{inter} values as in PS1, and the energy is compared with the corresponding FCI value. Here we consider optimal ($\mu_{\text{inter}}, \mu_{\text{intra}}, \nu, \nu'$) values, those giving CASSCF+V_{eff} energies $\leq$ 2 kcal/mol with respect to CIPSI energies. PS2 is designed to identify the evolution of the optimal parameters and their effect to the oveall PECs.

As observed in preliminary calculations for atomic and H_2 systems, each point in the PEC can be associated with multiple optimal parameters. This characteristic makes PS2 an exhaustive task, given the numerous parameters to be considered in each calculation. However, in an initial phase aimed at exploring the potential outcomes of varying the intra-term, we conduct PS2 by fixing μ_{inter} and restricting $\nu = \nu'$. The resulting optimal parameters indicate that variations in μ_{intra} primarily shift the total energy and do not play an essential role in describing dissociation. This result aligns with our previous observations in atomic and H_2 systems. Therefore, we decided to maintain the intra-atomic contribution (weighted by μ_{intra}) as a fixed and small parameter. Consequently, the results from PS2 consist of sets of optimal μ_{inter} (ν fixed) values for each point in the PECs.

6.6.1.1 Results from PS1

The isoelectronic diatomic molecules, N_2 and NO^+, have the same electronic configuration at the minima ($\sigma_g^2\pi_u^4\pi_g^{*0}\sigma_u^{*0}$) and both dissociate similarly into two atoms with spin quartets and 7 electrons on each atom. Particularly, the dissociation energy (D_e) of N_2 has often been used as a reference test of the capabilities of quantum chemical methods. This is because an accurate computation of D_e, comparable to the experimental value requires a very sophisticated treatment of the correlation problem. Some approaches can achieve a good agreement derived from cancellations of errors as well as from basis set superposition errors, as pointed out in previous studies [400, 402, 403]. However, D_e agreement with the experiment to within 1 kcal/mol is not even feasible within conventional multireference CI wave functions and large basis sets. Accurate estimations often requires accounting for core-valence correlation effects and including basis functions of angular momentum up to h. This problem is also present in more elaborate MRCI wave functions and extended ANO

basis sets with high angular momentum functions [400].

Motivated by these challenges, we explore our CASSCF+V_{eff} strategy for the calculation of corresponding PEC for N_2 and NO^+ singlet ground state. In Table 6.4 the most accurate results obtained from parametrizations based on the dissociation energy are depicted (respective PEC can be consulted in Appendix C, Figures 35 and 36). Optimum μ values lie in the range of -0.001 to 0.015, giving notable improvements in computed D_e's as illustrated by the reduction of percent errors from ∼12% (in the pristine CASSCF result) to ∼0.275% (with the optimum parameters), selected PEC displaying the lowest errors computed with $\nu = 0.5$. In fact, it is observed that the best CASSCF+V_{eff} outperforms NEVPT2 suggesting that the current approach might be more effective at introducing dynamic correlation than PT2-based approaches. Interestingly, for cases where the CASSCF+V_{eff} dissociation energy gives lowest percent errors, the magnitude of the multiplicative μ_{intra} parameters tend to be larger than μ_{inter}.

While for H_2 the changes induced by the effective potential to the Coulomb hole appear to be negligible (the V_{eff} contribution is inter atomic), in N_2 and NO^+ the effective potential has an impact on the Coulomb holes similar to that observed for atoms (where the contribution is intra-atomic, recall Fig. 6.2). Coulomb holes computed at equilibrium geometries are displayed in Figure 6.6. While CISD delivers a qualitative correct hole at short-range, CASSCF delivers a wrong description characterized by the positive values of the hole. Electronic distribution can be adjusted with the use of a positive intra-atomic. In cases were CASSCF exhibits erratic behavior at large inter electronic distances, the effective potential is incapable of correcting this. These findings indicate that the combination of parameters in the effective potential, involving a negative (small) *inter*-atomic term with a larger positive *intra*-atomic term enhances the accuracy of CASSCF energies. This combination delivers accurate D_e values while maintaining qualitatively correct electronic distributions. The *intra*-atomic contribution significantly influences the electronic distribution, specially at short-range, while the *inter*-atomic contribution plays a comparatively minor role in it.

Table 6.4: Effective potential parameters giving the lowest percentage of error in the computed dissociation energy with respect to FCI.

	N_2				NO^+		
ν	μ_{intra}	μ_{inter}	%Error[a]	ν	μ_{intra}	μ_{inter}	%Error[a]
0.5	0.001	-0.003	2.044	0.5	0.001	0.001	0.549
	0.001	-0.002	-2.372		0.001	0.002	-3.815
	-0.001	-0.003	2.044		0.001	0.003	-7.960
	-0.001	-0.002	-2.372		-0.001	0.003	-8.157
0.1	0.001	-0.002	4.985		-0.001	-0.002	0.133
	0.001	-0.001	-3.149		-0.001	0.002	-4.339
	-0.001	-0.002	4.985	0.1	0.001	0.001	-2.761
	-0.001	-0.001	-3.149		-0.001	0.001	-3.940
CASSCF(6,6)			-10.468	CASSCF(6,6)			-9.544
NEVPT2			-8.532	NEVPT2			-5.621

[a] Complementary data can be consulted in the Appendix C.

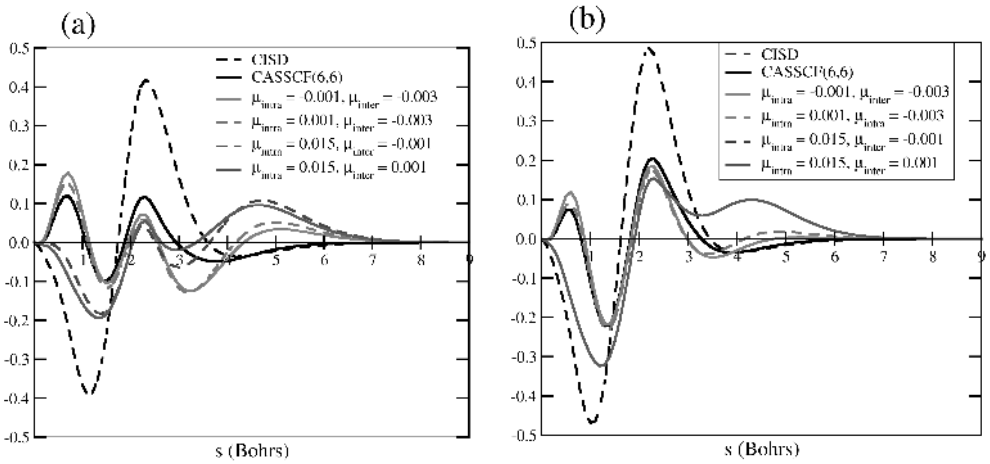


Figure 6.6: Computed Coulomb holes for (a) N_2 and (b) NO^+ diatomics at equilibrium bond lengths.

When extending the study to triplet ground state diatomics (O_2 and NH), differences with respect to singlet state molecules arise. In this case, accurate dissociation energies are obtained for cases where the magnitude of μ_{intra} is lower than μ_{inter}. This also, indicates that the *inter*-atomic contribution has a larger influence on the shape of PEC. Only in few cases the *intra*-atomic contribution is positive. These differences can be attributed to the nature of the ground state, since valence electrons in the triplet state experience a lesser degree of repulsion from each other, a larger energy correction can be put into the *inter*-atomic term. Due to electron separation, this term makes a negligible contribution to the energy at dissociation. However, near equilibrium, it must contribute significantly to the correction of D_e. In fact, as previously mentioned, the *intra*-atomic term only adds a shift to the total energy and has a smaller effect on the dissociation energy (corresponding PEC for these cases can be consulted in Figure 37 and 38, Appendix C).

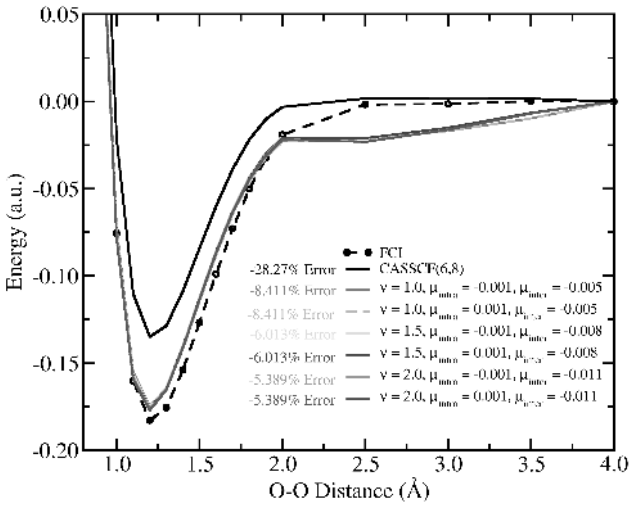


Figure 6.7: Ill-behaved PES ($\nu = 1.0, 1.5, 2.0$) for triplet ground state O_2, results computed with larger ν values show considerable distortions to the PES. Percentage of errors with respect to experimental dissociation values follows the same color code as that indicated by labels in the plot.

We must emphasize, the main problem with this parametrization scheme, is that only the

energies at dissociation and equilibrium are considered to select the optimal parameters. The use of constant μ_{intra} and μ_{inter} parameters along the entire PEC does not ensure that the resulting curve will behave correctly, hence PS1 can lead to artifacts in the PECs (see Figure 6.7, Appendix C). In this regard, point-by-point parametrization is more desirable. In the next section, results from the optimal fittings to FCI results along the PEC are discussed for all diatomics considered in this work.

Table 6.5: Effective potential parameters giving the lowest percentage of error in the computed dissociation energy with respect to FCI.

	O_2				NH		
ν	μ_{intra}	μ_{inter}	%Error[a]	ν	μ_{intra}	μ_{inter}	%Error[a]
0.5	-0.015	0.001	-0.151	0.5	0.001	-0.020	-0.640
	-0.015	-0.001	-0.151		0.001	-0.019	-2.576
	0.001	-0.003	-0.707		-0.001	-0.020	-0.640
	-0.001	-0.003	-0.707		-0.001	-0.019	-2.576
0.1	0.014	0.001	-1.461	0.1	0.001	-0.011	2.311
	0.015	0.001	1.050		0.001	-0.010	-1.421
	0.014	-0.001	-1.457		-0.001	-0.011	2.311
	0.015	-0.001	1.054		-0.001	-0.010	-1.421
CASSCF(6,8)			-25.720	CASSCF(4,4)			-32.552
NEVPT2			-3.140	NEVPT2			-20.223

[a] Complementary data can be consulted in the Appendix C.

6.6.1.2 Results from PS2

When applying PS2 to the dissociation of H_2, we observe that optimal μ_{inter} values (those for which the absolute error is $(E_{\text{CASSCF+}V_{\text{eff}}} - E_{\text{FCI}}) \leq 2.00$ kcal/mol), when plotted against inter-atomic separations, form a distinctive pattern with a specific shape (depicted in Fig. 6.8). This result allows us to identify a convenient selection of optimal μ_{inter} values by fitting the pattern to a function $\mu_{\text{inter}}(r_{AB})$ that rapidly converges to a constant value as $r_{AB} \to \infty$ and becomes finite for $r_{AB} \to 0$. Among the various possible functions satisfiying these requirements, the use of an error function is the most simple. This particular choice is notably suitable for H_2, as evidenced by the coefficient of determination ($R^2 > 0.99$, refer to Table 6.6) obtained when fitting the data to the function of the form:

$$\mu_{inter}(r_{AB}) = a(\text{erf}(b\, r_{AB}) - 1) \qquad r_{AB} \in [0, \infty) \tag{6.28}$$

where a and b are fitting parameters and r_{AB} is the inter-atomic distance. In the following, data fitting to function in Eq.6.28 will be referred as erf-fitting.

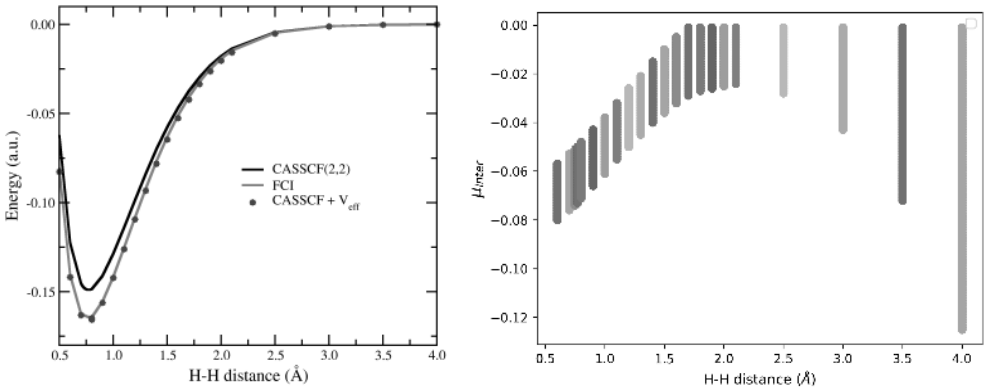


Figure 6.8: (left) Obtained H_2 PEC estimated from PS2 (computed with $\nu = 0.50$), the CASSCF+V_{eff} total energy is in perfect agreement with FCI/cc-pVTZ energy. (right) Corresponding optimum μ_{inter} values changing along the separation of H atoms, i.e. those for which CASSCF+V_{eff} energies are lower or equal to 2 kcal/mol.

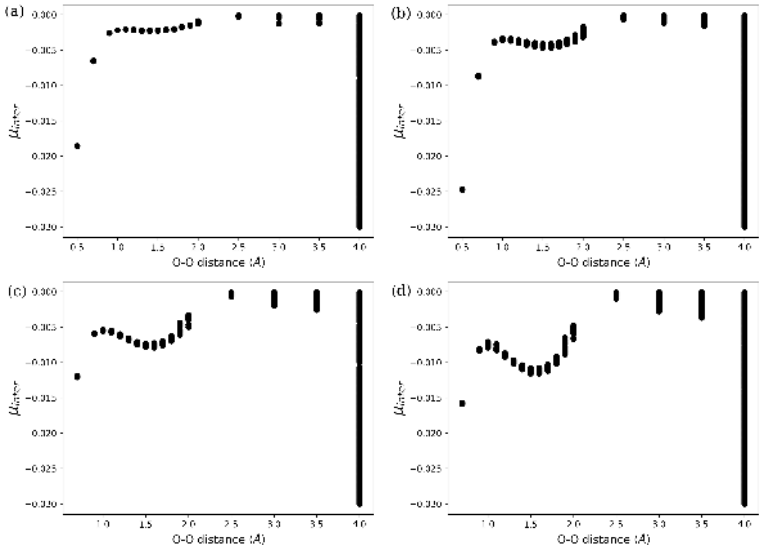


Figure 6.9: Optimum μ_{inter} values for O_2 along the PEC (computed with fixed $\mu_{intra} = 0.001$ value). Results computed with ν values: (a) 0.10, (b) 0.50, (c) 1.00, and (d) 1.50. A comparison of these figures reveals hints about the sensitivity of μ_{inter} values to ν parameter.

Figure 6.8 suggests a possible strategy to establish μ_{inter} values for other diatomics through erf-fitting. This idea is explored for the rest of diatomics, in such cases, the quality of erf-fitting changes due to deviations from the H_2 pattern-like results. μ_{inter} patterns display different shapes depending on the molecule, as illustrated in Figs. 6.9 and 6.10. Our exploration suggests that μ_{inter} patterns are sensitive to the choice of ν, becoming largely distorted as ν increases. This effect is specially enhanced at intermediate regions of the PEC. An illustrative example of this is depicted in Figure 6.9 for the O_2 molecule.

For heteroatomic dimers, larger pattern distortions appear compared to homoatomic dimers (as seen in the cases of N_2, NO^+, and NH in Figure 6.10). The erf-fitting features poorly ($R^2 < 0.7$) when applied to pristine μ_{inter} patterns. This difficulty is associated with these pattern changes. On the other hand, erf-fitting is appropriate for μ_{inter} points at both short and large interatomic distances, but it encounters challenges in accurately approximating

deviations at intermediary points in PECs.

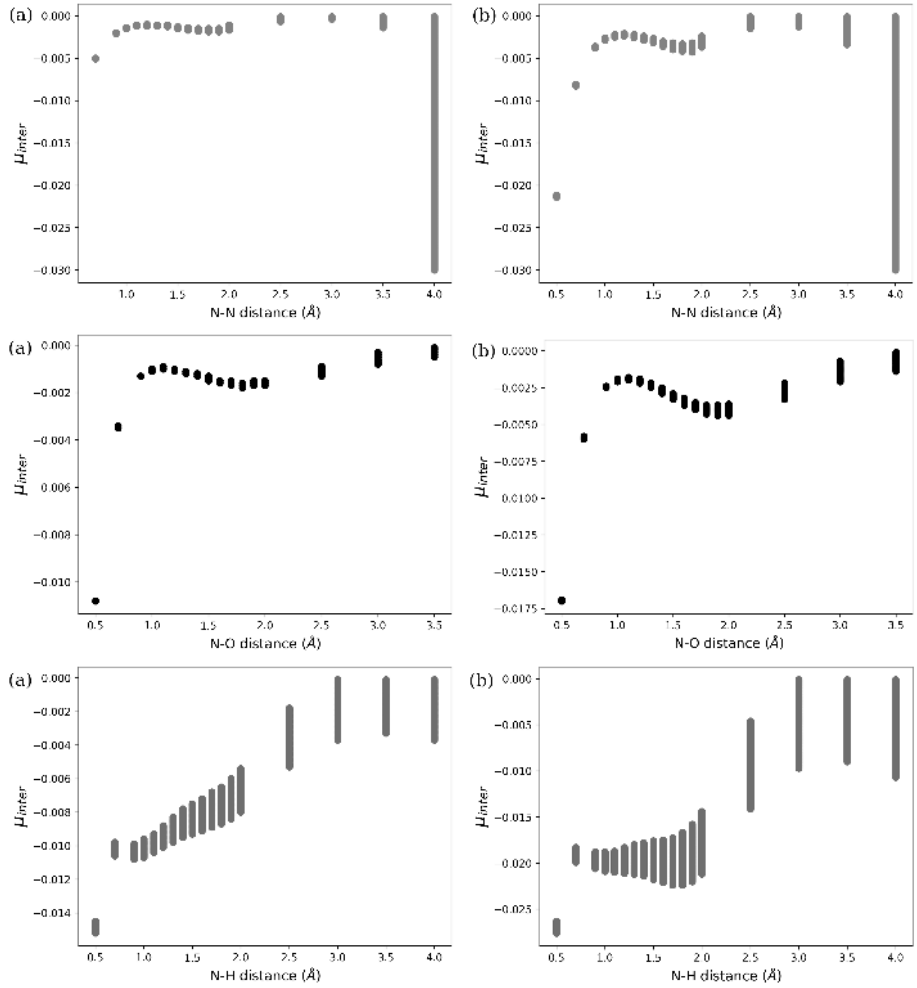


Figure 6.10: Illustration of μ_{inter} values which give CASSCF+V_{eff}/cc-pVDZ energies ≤ 2.00 kcal/mol with respect to FCI, along the PEC of N_2 (red), NO^+ (black) and NH (green). Results obtained with fixed $\mu_{intra} = 0.001$ value, and ν values: (a) 0.10, (b) 0.50. Note the pattern distortion increasing with the ν in comparing (a) with (b) cases, also note an enhanced effect present in heteroatomic dimers.

Achieving a correct fitting in such cases might necessitate different functions for each molecular system, a feature that is undesirable. Instead of persistently searching for the exact function, and guided by the results obtained in PS1, we opt to perform a reduced fitting approach. This involves eliminating values in intermediary regions of the PEC where distortion is significant and assuming that patterns approximately resemble the H_2 pattern. Ignoring intermediary PEC data ensures erf-fittings using Eq. 6.28 with $R^2 > 0.9$. With most of the data removed in the regions of 0.9-2.0 Å, this choice is a kind of extended version of PS1 adopted merely to explore the possibility of use erf-fitting strategy and study the role of μ_{inter}.

Tables 6.6 and 6.7 collect the numerical results obtained from such fittings. Despite data reduction, PECs obtained with fitted functions, show improved accuracy in the dissociation energies with respect to CASSCF. Particularly for H_2 PEC closely resembles the FCI result with percentage of errors of $\sim 0.07 - 0.08\%$. PEC for the rest of diatomics behave softly

along dissociation, showing good qualitative agreement with the FCI results (for cases giving the most accurate D_e, see Figure 6.11). As the values of μ_{intra} were selected to be positive, most of the PEC exhibit a shift to higher energies in comparison with prisitine CASSCF PECs (blue curves in Fig. 6.11).

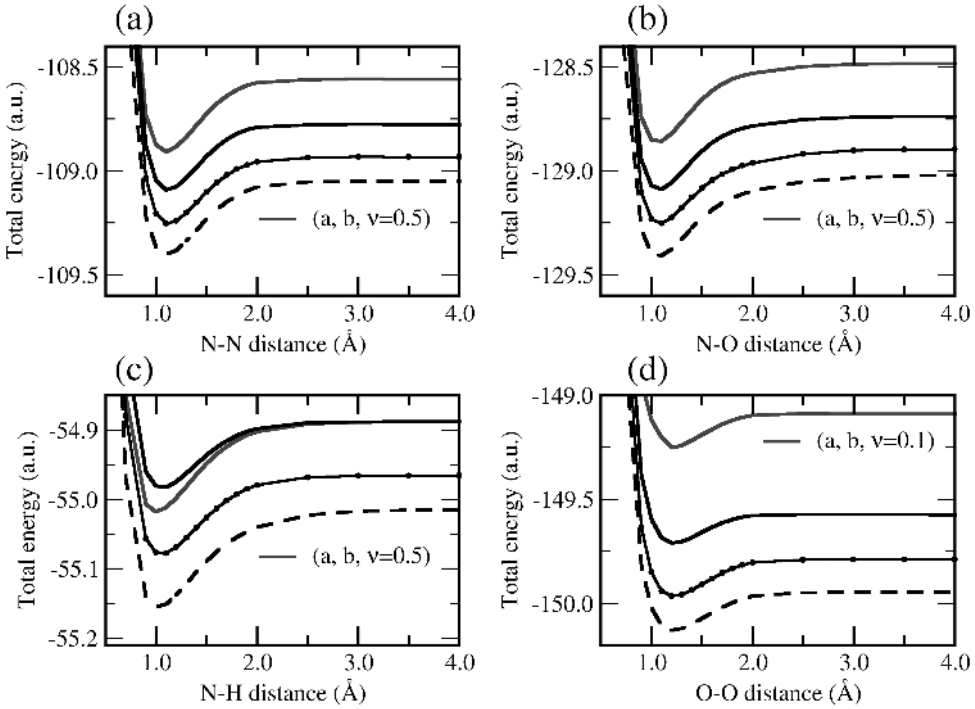


Figure 6.11: PECs delivering the most accurate D_e for (a) N_2 ($\mu_{\text{intra}} = 0.01$), (b) NO^+ ($\mu_{\text{intra}} = 0.01$), (c) NH ($\mu_{\text{intra}} = 0.00$) and (d) O_2 ($\mu_{\text{intra}} = 0.01$), computed with the erf-fitting strategy (blue continuous lines) a and b values used for the calculations can be consulted in Tables 6.6 and 6.7. The CASSCF/cc-pVTZ (black solid line), NEVPT2/cc-pVTZ correction (black solid line with dots), and CIPSI-FCI/cc-pVTZ (black dashed line) results are included for comparison.

In comparing the dissociation energies of diatomics, a significant improvement is achieved with the lowest error results. Generally, CASSCF+V_{eff} outperforms CASSCF, and the lowest error values are lower than the errors delivered by NEVPT2. When comparing CASSCF solutions with NEVPT2, the improvements are more significant for singlet state molecules. For instance, NEVPT2 corrections reduce the CASSCF percent error from -10.46% to -8.53% for N_2 (a reduction of 1.93%) and from -9.54% to -7.94% (a reduction of 1.60%) for NO^+ (see Table 6.6). However, this reduction is not as efficient for triplet state molecules, being 12.33% for NH and 22.58% for O_2 (see Table 6.7), with the latter showing larger corrections delivered by NEVPT2. Under the reduced fitting scheme, CASSCF+V_{eff} can correct CASSCF results, with the best corrections reducing the error by 9.64%, 6.23%, 25.19%, and 15.76% for the entire set of N_2, NO^+, NH, and O_2, respectively.

Table 6.6: Singlet state diatomics, resulting dissociation energies (in eV) and percentage of errors with respect to FCI from calculations with μ_{inter} fitted to the function $a(\text{erf}(br_{AB}-1))$.

Molecule	μ_{intra}	ν	a	b	R^2	D_e(eV)	% Error
H_2		0.1	0.0610	0.6493	0.9953	4.4980	0.1454
		0.1	0.0600	0.6093	0.9929	4.4980	0.1454
		0.5	0.1217	0.6115	0.9941	4.4950	0.0787
		0.5	0.1218	0.6095	0.9939	4.4950	0.0787
	CASSCF					4.05020	-9.7845
	FCI					4.4915	
N_2	0.000	0.1	0.0834	1.9469	0.9649	8.7090	-8.6876
	0.001	0.1	0.1328	2.3728	0.9900	8.6889	-8.8987
	0.010	0.1	0.1352	2.3796	0.9848	10.2660	7.6377
	0.000	0.5	0.0776	1.5614	0.9906	9.0443	-5.1726
	0.001	0.5	0.1123	1.9267	0.9469	8.7278	-8.4907
	0.010	0.5	0.1246	1.9709	0.9164	9.4592	-0.8217
	CASSCF					8.5392	-10.4679
	NEVPT2					8.7238	-8.5324
	FCI					9.5376	
NO^+	0.000	0.1	0.0788	2.169	0.9709	9.5860	-9.0538
	0.001	0.1	0.08987	2.267	0.9544	9.6532	-8.4163
	0.010	0.1	0.2842	3.0964	0.9668	10.8598	3.0312
	0.000	0.5	0.1376	2.2844	0.9255	9.5588	-9.3119
	0.001	0.5	0.1454	2.3203	0.9241	9.5947	-8.9713
	0.010	0.5	0.1194	2.1375	0.9671	10.1907	-3.3168
	CASSCF					9.5344	-9.5441
	NEVPT2					9.7034	-7.9404
	FCI					10.5403	

In comparison with PS1, CASSCF+V_{eff} dissociation energies worsen, but the errors remain lower than NEVPT2 results in most cases. While this can be considered a drawback in PS2 within the reduced fitting scheme, recalling previous results, it has been observed that in some cases, PS1 leads to distorted PECs, an undesirable feature that can be potentially fixed in full PS2. The use of an error function delivering μ_{inter} in the intermediary approach (PS2 with reduced data fitting) has shown to be an alternative, providing smooth PECs (see Fig. 6.11) and reducing the CASSCF dissociation energy error. While distortions were expected at intermediary regions in the PECs due to data reduction, no such apparent behavior was observed.

Table 6.7: Triplet state diatomics, resulting dissociation energies (in eV) and percentage of errors with respect to FCI from calculations with μ_{inter} fitted to the function $a(\text{erf}(br_{AB} - 1))$.

Molecule	μ_{intra}	ν	a	b	R^2	D_e(eV)	% Error
O_2	0.000	0.1	0.1043	1.9469	0.9649	3.7418	-24.2494
	0.001	0.1	0.1052	1.9627	0.9678	3.7631	-23.8198
	0.010	0.1	0.1011	2.0515	0.985	4.4474	-9.9653
	0.000	0.5	0.1041	1.6801	0.9901	3.8425	-22.2112
	0.001	0.5	0.1063	1.6968	0.9905	3.8425	-22.2112
	0.010	0.5	0.1188	1.8396	0.9968	4.0387	-18.2394
	CASSCF					3.6692	-25.7203
	NEVPT2					4.7846	-3.1400
	FCI					4.9397	
NH	0.000	0.1	0.0301	0.9542	0.9849	3.1192	-18.4244
	0.001	0.1	0.0294	0.9735	0.9834	3.1051	-18.7945
	0.010	0.1	0.0217	1.1404	0.9680	3.1040	-18.8229
	0.000	0.5	0.0343	0.4959	0.9762	3.5424	-7.3584
	0.001	0.5	0.0334	0.4811	0.9775	3.5644	14.8330
	0.010	0.5	0.0410	0.9410	0.9536	3.1154	-18.5241
	CASSCF					2.5790	-32.5521
	NEVPT2					3.0505	-20.2234
	FCI					3.8237	

In summary, PS2 strategy allow us to unveil optimum μ_{inter} values that can be fitted to an error function, suggesting a possible r_{AB} dependence for μ_{inter}. While this approach demonstrates success for H_2, its applicability diminishes in heteroatomic dimers, presenting challenges in achieving accurate fittings. The sensitivity of μ_{inter} patterns to the choice of ν, particularly in intermediate regions of the Potential Energy Curve (PEC), contributes to these challenges, therefore, in aggrement with our previous observations from results computed within the Roos' scheme, small ν values are also desirable when splitting V_{eff} into inter- and intra-atomic contributions. In general, obtained dissociation energies are worsened than those achieved in PS1, but the use of different values for μ_{inter} potentially improves the shape of PECs. Tables 6.6 and 6.7 summarizes the numerical outcomes of these fittings, providing a comprehensive overview of the enhanced accuracy achieved through this approach.

6.6.1.3 Analysis of selected systems in terms of correlation indices

Building on insights derived from the behavior of μ_{inter} and its potential dependence with interatomic separation, r_{AB}, this subsection conducts a comprehensive analysis of electron correlation along diatomic PECs. Motivated by complexities revealed in the previous parametrization strategies, our exploration focuses on NDC and DC, quantified by correlation indices $\overline{I}_{ND}$ and $\overline{I}_{D}$ respectively. The objective here is to identify a viable approach to incorporate these indices into V_{eff}, potentially facilitating the determination of μ_{intra} and μ_{inter}. Recently, $\overline{I}_{D}$ and $\overline{I}_{ND}$ have been proposed for the characterization of dynamic and nondynamic correlation, respectively [252]. These average quantities are computed from the natural orbital occupations as

$$\overline{I}_D = \frac{1}{2N}\sum_{i,\sigma}[\nu_i^\sigma(1-\nu_i^\sigma)]^{1/2} - \frac{1}{N}\sum_{i,\sigma}\nu_i^\sigma(1-\nu_i^\sigma) \tag{6.29}$$

$$\overline{I}_{ND} = \frac{1}{N}\sum_{i,\sigma}\nu_i^\sigma(1-\nu_i^\sigma), \tag{6.30}$$

where N stands for the number of electrons with spin σ, and ν_i^σ natural orbital occupation. For each point along the entire PEC, the natural orbital occupation numbers from the CASSCF solution were computed and used to calculate $\overline{I}_D$ and $\overline{I}_{ND}$, corresponding profiles are depicted in Fig. 6.12.

The behavior of $\overline{I}_{ND}$ can be described as follows, for $r_{AB} \to 0$, $\overline{I}_{ND} \to 0$ and progressively increases up to a constant value at dissociation. This behavior is similar to that observer for μ_{inter} patterns along H_2 dissociation. While $\overline{I}_{ND}$ behaves similarly for all diatomics under study, $\overline{I}_D$ shows a different behavior depending on the spin state of the molecule. For singlet state molecules (N_2 and NO^+), $\overline{I}_D$ is larger than zero when $r_{AB} \to 0$ and increases up to a maximum near bonding regions in the PES ($\sim$1.0-2.0 Å, where μ_{intra} pattern distortions are present), and eventually decays to zero at dissociation, see Fig. 6.12 (a) and (b). In triplet state diatomics (O_2 and NH), $\overline{I}_D$ behaves similarly at short and bond regions and then acquires a constant value at dissociation, see Fig. 6.12 (c) and (d). Interestingly, the rise of $\overline{I}_D$ at bond and dissociation regions in PECs coincides for the cases where the optimum μ_{inter} patterns show larger distortions as previously exposed.

These observations have led us to propose the use of correlation indices as scaling factors in the components of the effective potential. Since they exhibit a behavior quite similar to the patterns of μ_{inter} previously presented. However, we lack evidence that attests to a direct relationship between these patterns and the correlation indices. On the other hand, the use of such indices could be desirable, as the dependence on r_{AB} would be decoded into a dependence on natural orbitals (ν_i). Therefore, we propose to explore a data fitting strategy involving the use of $\overline{I}_{ND}$ and $\overline{I}_D$ to approximate the multiplying (μ) factors of the effective potential.

The previous study on atoms highlighted the sensitivity of the CASSCF+V_{eff} energy to small changes in μ, leading to deviations from accurate FCI results. Additionally, it was asserted that the intra-atomic term in V_{eff} adequately characterizes atomic systems, as they lack inter-atomic contributions. Consequently, it is anticipated that intra-atomic terms in diatomics should remain reasonably constant throughout dissociation. Therefore we propose the following form of the effective potential depending on correlation indices

$$V_{\text{eff}} = Y\overline{I}_{ND}\sum_{i<j}\frac{e^{-\nu r_{ij}^2}}{r_{ij}} + X\overline{I}_D\sum_{p<q}\frac{e^{-\nu' r_{pq}^2}}{r_{pq}}, \tag{6.31}$$

where X is a constant parametrized to ensure that the magnitude $X\overline{I}_D$ becomes of the same order as the previously determined μ_{Inter} values yielding energies close to FCI (as determined in PS2). And Y is a constant, parametrized to ensure that the $Y\overline{I}_{ND}$ product becomes of the order (10^{-3} in most cases) of the previously explored μ value for atoms giving correlation energies of FCI quality. Therefore, we relate the term containing $X\overline{I}_D$ to the inter-atomic contribution, an the term containing $Y\overline{I}_{ND}$ to the intra-atomic one.

Dissociation energies computed under this fitting scheme, and their respective errors with respect to FCI, for different selection of the parameters can be consulted in Table 6.8.

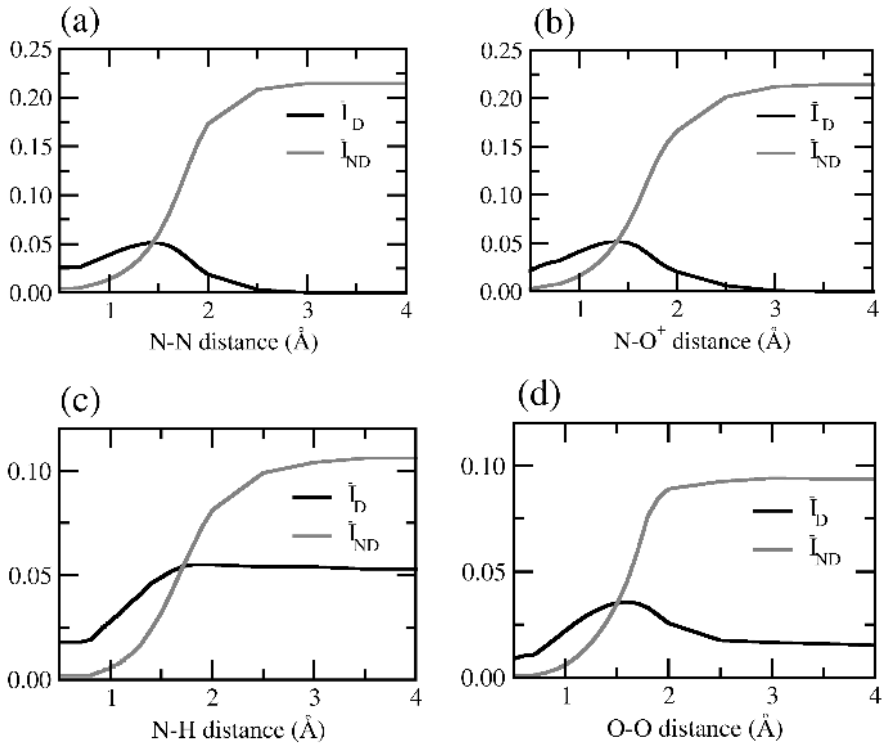


Figure 6.12: Correlation indexes $\overline{I}_D$ and $\overline{I}_{ND}$ computed at CASSCF/cc-pVDZ level, for the four diatomics (a) N_2, (b) NO^+, (c) NH, and (d) O_2 along the dissociation coordinate.

While systematic improvements in dissociation energies are not readily apparent, consistent patterns emerge across both the current and previous erf-fitting schemes. Singlet state molecules exhibit the lowest percentage errors, showcasing the efficacy of the optimized potential in introducing dynamic correlation and enabling dissociation energy calculations close to FCI results. Notably, for N_2 and NO^+, the percentage errors are as small as 1.90% and 1.15%, respectively. However, despite these successes, dissociation energies are frequently overestimated, resulting in positive percentage errors, especially evident for triplet state molecules, with NH showing the largest errors. The improvements in CASSCF energies for triplet state molecules are limited, leading to the highest percent errors.

A crucial insight is that μ_{inter} exhibits molecule-dependent and, more importantly, r_{12}-dependent behavior. Although the elimination of r_{12} dependency through the use of $\overline{I}_D$ has been achieved, the absence of a universal Y value capable of providing accurate results suggests that molecule-dependency is not fully captured in the current approach.

In comparing with NEVPT2 calculations, as before, it is seen that optimum results yield dissociation energies whose values lie in between the CASSCF and NEVPT2 result. Comparisons of corresponding total energy PECs are depicted in Figure 6.13 (red curves). For singlet state molecules, obtained PEC display a soft behavior along the dissociation, while for triplets, this is not completely achieved, since PECs can be altered at intermeriary regions in a similar fashion as previouly shown in PS1 (see Fig. 6.13 (d)). The latter effect could suggest a disproportion in the correlation treatment caused by the non-zero $\overline{I}_D$ values at

long-range separations. Alternatively, this could indicate that the inter-term should decay faster, a condition achieved with larger ν. In fact, results for $\nu = 0.5$ generally appear quite satisfactory, as depicted in Table 6.8 (except for NH).

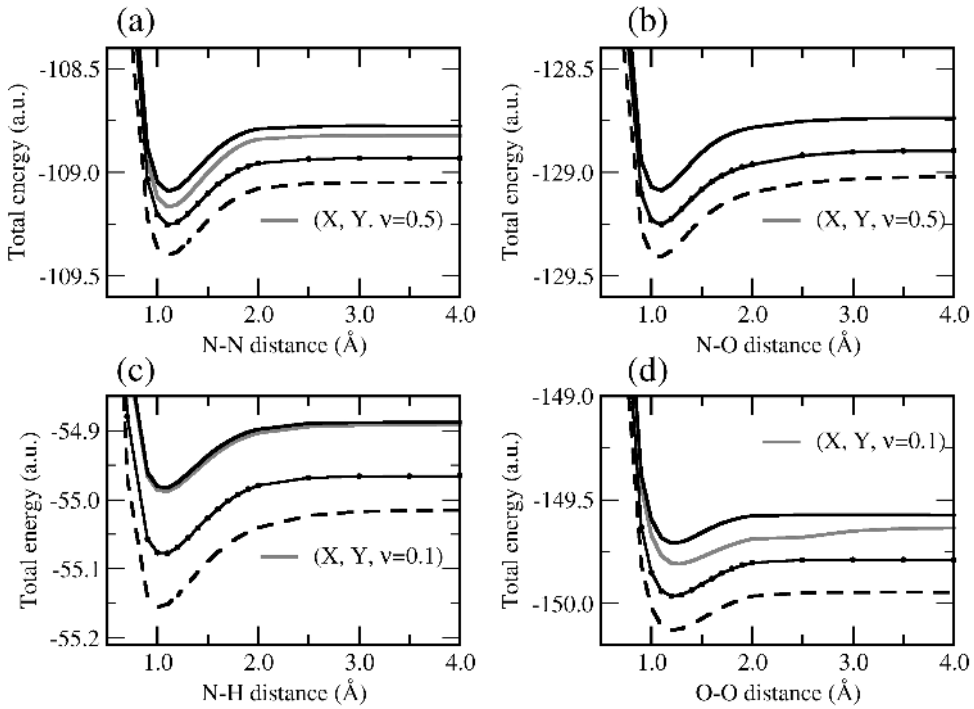


Figure 6.13: PECs delivering the most accurate D_e for (a) N_2, (b) NO^+, (c) NH and (d) O_2, computed with correlation indices fitting (red continuous line). Optimum fitting X, Y values used for the calculation can be consulted in Table 6.8. The CASSCF/cc-pVTZ (black solid line), NEVPT2/cc-pVTZ correction (black solid line with dots), and CIPSI-FCI/cc-pVTZ (black dashed line) results are included for comparison.

The exploration presented here serves merely to identify a possible route to encode information from correlation indices into V_{eff}. The work presented here represents an initial step in such a direction. While not reaching the proximity of FCI results (as in some of the previous schemes), the form of the effective potential proposed here appears to yield sufficiently improved results with respect to CASSCF energies. While the effective potential based on correlation indices is always feasible, it does not always yield good results. However, when applied to singlet state molecules, the results appear to be more favorable compared to triplet state molecules. Additionally, we note that the difference in atomic volume (not shown here) may be a matter of relevance, as in the case of NH, additional calculations related with QTAIM quantities provide some support to this observation (See Appendix C).

Table 6.8: Resulting dissociation energies and percentage of errors with respect to FCI from calculations optimizing $\mu_{inter} = -0.1 I_D$ and using different choices for $\mu_{intra} = Y I_{ND}$.

Molecule	Y	ν	D_e(eV)	% Error
N2	0.00	0.1	12.8021	34.2282
	-0.001	0.1	12.6231	32.3509
	-0.01	0.1	11.0125	15.4636
	0.01	0.1	14.5867	52.9387
	0.00	0.5	10.6301	11.4551
	-0.001	0.5	10.4190	9.2411
	-0.01	0.5	9.3561	-1.9030
CASSCF			8.5392	-10.4679
NEVPT2			8.7238	-8.5324
FCI			9.5376	
NO+	-0.001	0.5	11.7129	11.1243
	-0.01	0.5	10.4191	-1.1499
CASSCF			9.5344	-9.5441
NEVPT2			9.7034	-7.9404
FCI			10.5403	
O2	-0.001	0.1	5.6246	13.8655
	-0.01	0.1	4.6844	-5.1672
	-0.001	0.5	4.6428	-6.0100
	-0.01	0.5	3.9995	-19.0327
CASSCF			3.6692	-25.7203
NEVPT2			4.7846	-3.1400
FCI			4.9397	
NH	-0.001	0.1	2.7375	-28.4088
	-0.01	0.1	2.0128	-47.3598
	-0.001	0.5	2.6447	-30.8354
	-0.01	0.5	2.1641	-43.4031
CASSCF			2.5790	-32.5521
NEVPT2			3.0505	-20.2234
FCI			3.8237	

6.7 Conclusions

In this work, we revisit the use of an effective potential accounting for the missing dynamic electron correlation from a CASSCF wave function. Preliminary results on the calculation of correlation energies for small atomic systems manifest the potentially high accuracy achieved within this method. Provided a proper parametrization, correlation energies of FCI quality can be determined. Nevertheless, the inherent arbitrariness in the parametrization process poses a challenge in extending the method's applicability. Indeed, establishing a robust and practical parametrization scheme for this type of effective potential appears to be a significant challenge.

Nevertheless, we have identified the consequences of the effective potential on both the

energy and the electronic distribution of the system. In atoms, a negative (attractive) potential leads to a substantial reduction in the total energy, however, this results in discrepancies in the electronic distribution, as depicted by the corresponding Coulomb holes. Within a positive (repulsive) potential, the effect is the opposite, leading to increments in the energy with improvements in the electronic distribution at short-range.

For the case of H_2, the parametrization of the effective potential shows that, irrespective of the ν value, μ_{inter} values near 0.020 consistently yielded the lowest errors in dissociation energies compared to FCI/cc-pVDZ reference values. The effective potential reduces the percent error in the CASSCF dissociation energy from $\sim 6\%$ to $\sim 0.6\%$. While PECs (and energies) can be improved by a suitable selection of the parameters (μ_{inter}, ν), Coulomb holes are poorly affected regardless of the sign of μ_{inter}, suggesting that the effective potential has a negligible effect on the Coulomb hole.

Motivated by these results and by the goal to achieve accurate description of PECs of diatomic molecules with atoms heavier than H, we propose to separate the effective potential into inter- and intra-atomic terms. Our numerical data obtained by parametrization schemes based on dissociation energies (PS1) or by adjustment of the total energy point-by-point in the PEC (PS2), these allow us to remark the following effects, trough our exploration of the results delivered by different selection of parameters: we identify the intra-atomic term as responsible of large effects on the electronic distribution at short-range (mostly observed in atomic systems), characterized by the qualitative behavior of the corresponding CASSCF+V_{eff} Coulomb holes. Our experience shows that, in order to get a balanced improvement of CASSCF energy and Coulomb hole, the intra-atomic terms must be positive and lie in values of the order of 0.01 (as much). For diatomics, the intra-atomic contribution merely shifts the energy profile, but its presence in the effective potential is needed to correct the Coulomb hole shape. Regarding the inter-atomic term (not present in atoms), its relevance impacts more on dissociation energies and the shape of the PECs.

Two additional strategies to fit data obtained from PS2 were taken into account. The first involves a relation of the magnitude of the inter-atomic contribution to the atomic separation r_{AB} through the function $\mu_{\text{inter}} = a(\text{erfc}(br_{AB}) - 1)$. In this regard, PS2 data fitting is challenging for intermediate r_{AB} values, so we consider only points at short and large inter-atomic distances.

The second fitting strategy was made in terms of average correlation indexes $\overline{I}D$ and $\overline{I}ND$, associated with each point in PECs. This selection was an attempt to transfer the r_{AB} dependency to $\overline{I}D$ and $\overline{I}ND$. However, corresponding results do not show systematic improvements that can be generalized to establish an analytic form for the μ_{inter} parameter, but they allow us to identify further perspectives for the development of the CASSCF+V_{eff} approach. In fact, it seems that the fitted parameters (X and Y) are not universal, i.e., they are molecule-dependent. We need a more exhaustive analysis with variants of $\overline{I}D$ and $\overline{I}ND$ that are not size-intensive. Otherwise, investigate how the parameters depend on features of the molecule (e.g., atomic volumes, atomic numbers, etc.).

Regardind ν exponential parameter, in atoms, optimal values can be related to μ by a simple linear relation, while for diatomics, ν has effects on both, the total energy and electronic distribution that are mainly manifested in μ_{inter} . From results obtained in PS2, it is observed a large variation of optimal μ_{inter} values as ν grows. Values in the range $\nu = 0.10$ to 0.50 seem to give satisfactory results in general, while larger values must be avoided since potentially deliver too much correlation energy into calculations.

These findings provide valuable insights into the parametrization of the effective potential and its impact on atomic and diatomic systems, offering a basis for further research and refinement of this approach. Further advances pursuing for a refinement in the parametrization scheme, elucitdation of a suitable form of the effective potential, and a desirable extention to larger systems, are in current development in our laboratory.

CHAPTER 7

Spin-Orbit Couplings of Open-Shell systems with Restricted Active Space Configuration Interaction

7.1 Abstract

In this Chapter we perform electronic structure calculations to unravel the origin of SOCs in open-shell molecules. For that, we select systems displaying di or polyradical character, e.g., trimethylene, analyze the changes in the magnitude of SOC constants along molecular distortions of ethylene, and in the presence of intermolecular interactions between open and closed-shell moieties in the $O_2 - C_2H_4$ system. Calculations were performed by using non-relativistic wave functions obtained with the RASCI method, in conjunction with a recent implementation of the calculation of SOC based on the spin-orbit mean field approximation. Our results demonstrate the suitability of RASCI in the calculation of SOCs of open-shell systems, while providing a deep understanding of the relationship between couplings and the nature of the electronic states. Moreover, we introduce a new definition of the SOC constant for the study of molecular aggregates.

7.2 Introduction

The non-relativistic picture assumes that Coulomb interactions between charged particles are instantaneous. Consequently, the interaction between electronic states of different spin is neglected [434]. The mixing of spin-pure states can be achieved by introducing relativistic effects such as spin-orbit and spin-spin interactions. Concretely, electronic SOC arises from the interaction between the angular momentum of the electron with the intrinsic magnetic moment associated with its spin. The presence of efficient SOC can have a crucial impact on the fate of chemical and photophysical reactions, enabling for non-radiative, e.g., intersystem crossing (ISC) [435], and radiative, e.g., phosphorescence [436], processes to take place, or tuning the profile of X-ray absorption spectra [437]. Although SOC is specially intense in the presence of heavy elements, it has also been shown to play a decisive role in shaping the properties of molecules with light elements [438, 439, 435].

The effect of SOC can be also present and important in open-shell systems. For instance, SOC is fundamentally relevant in reactions involving diradical formation, as in the reaction of molecular oxygen with unsaturated hydrocarbons [440, 441, 442, 443, 444, 445, 446, 447, 448, 449], sensitization processes involving singlet oxygen [439], reaction pathways involving transition metals [450], and thermal generation of triplet states from dioxetanes [451]. Similarly, accounting for SOC might be crucial in order to obtain high-accuracy chemical bond dissociation energies, since SOC tends to affect the most the energy of the open-shell products [452].

From a theoretical point of view, the early work of Salem and Rowland [275] tried to rationalize the main aspects controlling the magnitude of singlet-triplet SOC in diradicals. From their analysis, the authors identified three general features controlling SOC in diradicals: (i) orientation of the two singly occupied atomic orbitals. Concretely, strong SOC requires the orbital axes to be orthogonal to each other and to the axis around which orbital angular momentum is being created. (ii) SOC is proportional to the degree of ionic character of the singlet state. If the singlet is purely covalent, with one electron on each atomic orbital, the triplet-singlet SOC vanishes. (iii) SOC decreases sharply as the distance between radical centers increases. However, too close fragments can imply a poorer orbital orthogonality with a suppression of SOC. In general, a favorable overlap is needed to maximize SOC.

The principal objective of this Chapter is to evaluate and rationalize SOC open-shell systems. We aim to establish a relationship between SOCs and the electronic structure details of the interacting states, and the structural dependency of the computed couplings. To this end, we employ non-relativistic wave functions described by the RASCI method [380, 115], together with a recent implementation that allows the computation of SOC by means of the Breit-Pauli Hamiltonian [34, 453]. The method makes use of the spin-orbit mean field approximation and the spinless transition density matrix [268].

The content of this Chapter is organized as follows. First, we review the theoretical aspects of new implementations devoted to SOCs calculations within RASCI. Then, we provide the essential computational details of the study, and discuss the obtained results in the application of our methodology in several molecular systems. Finally the conclusions section summarizes the main findings of this work.

7.3 Theory

7.3.1 The Breit-Pauli Hamiltonian

Relativistic effects derived from the spin-orbit interaction can be evaluated with non-relativistic ansätze through perturbation theory. For that, it is standard practice to employ the spin-orbit Breit-Pauli (BP) [34, 453] Hamiltonian, valid for weak SOCs and expressed as:

$$H_{SO}^{BP} = \frac{1}{2c^3}\left[\sum_i \mathbf{h}^{SO}(i)\cdot\mathbf{s}(i) + \sum_{i\neq j}\mathbf{h}^{SOO}(i,j)\cdot(\mathbf{s}(i)+2\mathbf{s}(j))\right], \tag{7.1}$$

where c is the speed of light, $\mathbf{s}(i)$ is the spin of electron i, and $\mathbf{h}^{SO}(i)$ and $\mathbf{h}^{SOO}(i,j)$ are one- and two-electron operators defined as:

$$\mathbf{h}^{SO}(i) = \sum_I \frac{Z_I}{\mathbf{r}_{iI}^3}(\mathbf{r}_{iI} \times \mathbf{p}_i), \tag{7.2}$$

$$\mathbf{h}^{SOO}(i,j) = -\frac{1}{\mathbf{r}_{ij}^3}(\mathbf{r}_{ij} \times \mathbf{p}_i), \tag{7.3}$$

with Z_I denoting the bare nuclear charge of nucleus I, $\mathbf{p}_i$ represents the momentum of electron i, $\mathbf{r}_{iI}$ the relative coordinates of electron i and nucleus I, and $\mathbf{r}_{ij}$ that between electrons i and j. The matrix element of the BP operator between two Slater determinants differing by the occupation of one spin orbital ($\phi_a \leftrightarrow \phi_b$) is given by:

$$\begin{aligned} H_{ab}^{BP} = & \langle\phi_a|\mathbf{h}^{SO}(1)|\phi_b\rangle + \frac{1}{2}\sum_k n_k(\langle\phi_a\phi_k|\mathbf{h}^{SOO}(1,2)|\phi_b\phi_k\rangle \\ & - \langle\phi_a\phi_k|\mathbf{h}^{SOO}(1,2)|\phi_k\phi_b\rangle - \langle\phi_k\phi_a|\mathbf{h}^{SOO}(1,2)|\phi_b\phi_k\rangle), \end{aligned} \tag{7.4}$$

where n_k denotes the occupation of the orbital *k*. It is worth noting that, from an independent-particle model point of view, eq 7.4 defines a Fock-type operator for a single determinant approximation, i.e., it describes the mean spin-orbit interaction of an electron in a field generated by all the electrons in $\{\phi_k\}$ with $n_k = 1$. Eq 7.4 has inspired a spin-orbit mean-field (SOMF) approximation [454] for the two-electron part of the BP Hamiltonian (eq 7.3). In second-quantization form, it can be written as:

$$\begin{aligned} H^{SOMF} &= \sum_{pq}\left[I_{pq} + \sum_{rs}\rho_{rs}(J_{prqs} - J_{prsq} - J_{rpqs})\right]a_p^\dagger a_q \\ &= \sum_{pq}\left[h_{pq}^{1eSO} + h_{pq}^{2eSOMF}\right]a_p^\dagger a_q \end{aligned} \tag{7.5}$$

where h_{pq}^{1eSO} and h_{pq}^{2eSOMF} are the spin-orbit one and two-electron mean-field matrices, ρ_{rs} is a density matrix, and I_{pq} and J_{prqs} respectively are one- and two-electron spin-orbit integrals,

$$I_{pq} = \langle\phi_p(1)|\mathbf{h}^{SO}(1)\cdot\mathbf{s}(1)|\phi_q(1)\rangle \tag{7.6}$$

$$J_{pqrs} = \langle\phi_p(1)\phi_q(2)|\mathbf{h}^{SOO}(1,2)\cdot(\mathbf{s}(1)+2\mathbf{s}(2))|\phi_r(1)\phi_s(2)\rangle \tag{7.7}$$

Application of the Wigner-Eckart theorem allows to express eq 7.5 as the contraction of the spinless transition matrix (u in eq 7.8) with one- and two-electron mean-field terms scaled by Clebsch-Gordan coefficients, as described by Pokhilko et. al. in the frame of equation-of-motion coupled cluster (EOM-CC) [455, 456],

$$u_{pq} = \frac{1}{\sqrt{2}}(\gamma_{pq} - \gamma_{\bar{p}\bar{q}})/\langle S'M;10|SM\rangle \tag{7.8}$$

$$\gamma_{pq} = \langle ISM|a_p^\dagger a_q|I'S'M'\rangle \tag{7.9}$$

where γ_{pq} is a matrix element of the transition density between states I and I' respectively with spins S and S' and spin projections M and M', and overbar indices in $\gamma_{\bar{p}\bar{q}}$ indicate spin-β orbitals. The $\langle S'M;10|SM\rangle$ factor in eq 7.8 is Clebsch-Gordan coefficient.

While SOCs of multiple components depend on the orientation in space, the spin-orbit coupling constant (SOCC) is rotationally invariant. A great advantage of the described algorithm is the possibility to compute the entire SOC matrix for all spin projections of the two electronic states, which are necessary to calculate the SOCC, from the calculation of a single pair of microstates [457, 458]. In molecules, the SOCC is obtained by taking the sum over all spin projections as follows:

$$|SOCC|^2 = \sum_{M,M'} \left| \left\langle SM \middle| H^{SO} \middle| S'M' \right\rangle \right|^2 \tag{7.10}$$

7.4 Computational details

Electronic states have been computed with the RASCI approach including up to hole and particle excitations (unless indicated). SOC calculations were performed considering the recommendations from a previous work [268], as the role of the basis set dependence, balanced active spaces (unbalanced spaces deteriorates the SOC results), and density symmetrization (which does not have a large impact on the couplings). It is important to notice that the hole and particle truncation of the excitation operator is, in general, unable to recover most of the (dynamic) electron correlation [115]. Whereas the lack of effective dynamic correlation can be notably detrimental for the accuracy of the computed relative energies, it is expected to have a lesser impact on SOCs [268]. Dynamic correlation can be related to the sum of small contributions from excited configurations, which can notably affect electronic energies, while having a lesser impact on the electronic structure composition. Hence, couplings, which depend on the nature of electronic states, are expected to be less effected than electronic state energies. On the other hand, strong or static correlations involve large mixing of different configurations, which might notably tune SOCs.

In order to provide a deeper picture of the mechanism of the SOC interaction, we perform the calculation of natural transition orbitals (NTOs) of the spinless transition density matrix (eq 7.8) and their participation ratios as prescribed by Krylov [256]. Here is important to mention that, we follow Q-Chem's standard molecular orientation, which is different from that of Mulliken, thus symmetry labels for states can be different from those placed in literature. In cases of interest, we compute the Yamaguchi [239] (N_Y) and Head-Gordon [240] (N_U) indexes (Eqs. 2.163 and 2.164). Spatial localization of unpaired electrons has been characterize by the representation of the FOD [115]. All calculations in this Chapter have been carried out with a development version of the Q-Chem package [381].

7.5 Results and discussion

In this section we compute spin-orbit interactions in the trimethylene molecule to investigate the capabilities of our methodology applied to biradicals; we explore different molecular distortions of the ethylene molecule to analyze how the interconversion between closed- to open-shell structure affects the computed couplings; and we consider the ethylene-oxygen dimer to understand how intermolecular interactions tune SOCs.

7.5.1 Trimethylene biradical

The trimethylene molecule (TM) is a paramount example of an organic biradical, which has been largely explored both experimentally and computationally [459, 460, 461, 462, 463, 464, 465, 466]. Experimentally, it has been observed that TM can be driven towards an heterogeneous reaction (isomerization) to give propylene, or by stereoisomerization it can form cyclopropane. Common reactions of TM result in a mixture of these products [459], with a propylene/cyclopropane ratio that can be controlled by the electronic nature of the precursor[460]. This behavior motivated to study the differences in chemical reactivity with the spin multiplicity (singlet or triplet) of the biradical and its precursors [467, 461, 463], and the interconversion between the triplet and singlet manifolds through ISC [462]. Furlani and King [464] calculated SOCs at the MCSCF level in the vicinity of the singlet-triplet intersection, whereas Carlacci and co-workers [468] used the same *ab initio* procedure to examine the molecular geometry dependence of SOCs. Motivated by these studies, here we test RAS-SF wave functions to evaluate singlet-triplet SOCs for different structural conformers of the TM biradical. Molecular structural parameters (bond lengths and angles) have been taken from the work of Doubleday and co-workers [462]. Studied molecular geometries correspond to the torsion of edge CH_2 groups defined by the two HCCC dihedral angles α and β as the counterclockwise rotations from the CCC reference plane (Figure 7.1). Accordingly, structural conformers will be labeled as (α, β). Structures (0, 0) and (90, 90) belong to the C_{2v} symmetry point group, with $^3B_1$ and $^1A_1$ being the lowest triplet and singlet states, respectively. The (0, 90) conformer present a C_s symmetry, with lowest lying $^3A''$ and $^1A''$ states, and those structures with $\alpha = \beta \neq 0°,\ 90°$ belong to the C_2 symmetry.

The relative torsion of the CH_2 side groups modifies the orientation of the carbon $2p$ orbitals and has a large impact on the biradical character of the system. Simultaneously, torsion induced changes in the frontier molecular orbitals tune the nature of the lowest singlet and triplet states, causing sizeable variations on the computed SOCs (Table 7.1). In the following, we explore singlet-triplet RAS-SF SOCs for the different TM conformers, the dependence of the computed couplings with the basis set, the size of the fully correlated space (RAS2), and the truncation level of the (spin-flip) excitation operator. In all calculations, we employ the lowest energy restricted Hartree-Fock triplet determinant as the reference configuration (Ψ_0 in eq 2.59).

SOCC virtually vanishes at the (0, 0) configuration, which has been related to the weak overlap between the frontier orbitals, i.e., those defining the two unpaired electrons, in agreement with Furlani and King results [464]. Molecular torsions inducing larger orbital overlaps tend to result in larger couplings. The strongest SOCC is obtained for the (90, 90) structure, for which the overlap between the singly occupied $2p$ orbitals is much efficient. At the (0, 90) conformer, the relative orbital orientation is ideal, but SOC vanishes due to the poor ionic character of the singlet wave function. In general, the magnitude of SOCCs correlates to the biradical nature of the singlet state as quantified by N_U, with larger couplings associated to weaker diradical character, that is, smaller N_U values. These results can be related to the rationalization done by Furlani and King in terms of the ionic character of the singlet state, since ionic configurations imply the pairing of electrons.

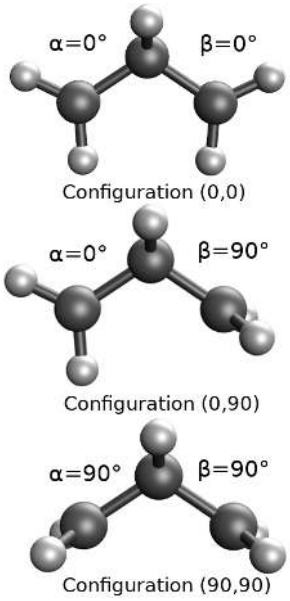


Figure 7.1: Representation of three investigated structural configurations of the TM biradical: (0,0) and (90,90) with C_{2v} symmetry, and the C_s structure (0, 90).

Table 7.1: RAS-SF SOCCs (in cm^{-1}) computed with (6e, 5o) RAS2 space and the cc-pVXZ (X = D, T, and Q) basis set family. Values for the Head-Gordon N_U index for the singlet state and vertical singlet-triplet gap ($\Delta E_{ST} = E_S - E_T$ in meV) computed at the RAS-SF/cc-pVQZ level.

	SOCC				
(α, β)	cc-pVDZ	cc-pVTZ	cc-pVQZ	N_U	ΔE_{ST}
(0, 0)	0.007	0.005	0.005	1.88	37
(0, 90)	0.056	0.058	0.058	2.01	80
(35, 35)	0.356	0.371	0.375	1.54	-12
(145, 35)	0.439	0.493	0.515	1.74	45
(150, 60)	0.802	0.877	0.906	1.60	22
(45, 60)	1.242	1.315	1.340	1.30	-105
(35, 90)	1.276	1.367	1.400	1.39	-49
(90, 90)	3.244	3.456	3.536	1.01	-291

The magnitude of SOCCs can be further analyzed by the representation of the natural transition orbital (NTO) pairs of the spinless triplet transition density matrix (u in eq 7.8). In Figure 7.2, we show the two leading spinless NTO pairs, i.e., those with the largest weights (ω), related to the singlet-triplet SOC for conformers (0, 90) and (90, 90). The transition in the (0,90) configuration is built from two degenerated contributions that correspond to the spin-flip excitation within the same p-like orbital, mostly localized on one of the edge carbon atoms. These terms do not involve change in orbital character needed for sizeable couplings (El-Sayed's rule [469]), which explains the nearly zero SOCC. On the other hand, the two main hole/electron terms for the (90, 90) case imply a change in orbital angular momentum, resulting in non-vanishing spin-orbit interactions, almost two orders of magnitude larger than in (0, 90). RAS-SF SOCCs show a basis set dependence along the cc-pVXZ (X = D, T, Q) family (Table 7.1), with couplings increasing with the size of the basis set (except for the (0,

0) conformer) and variations from cc-pVQZ values in general no larger than 10% and 5% for cc-pVDZ and cc-pVTZ, respectively.

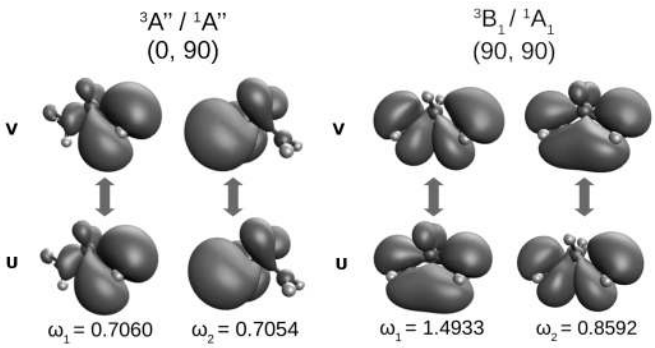


Figure 7.2: Dominant NTO pairs (isovalues of 0.05) of the spinless transition matrix of TM involved in the singlet-triplet SOC at (0, 90) and (90, 90) torsion angle configurations. The respective singular values ω_1 and ω_2 are depicted below each pair.

Next, to investigate the role of hole and particle contributions to the spin-orbit interactions, we use a RAS2 space with two electrons in two orbitals (2e, 2o) and the 3-21G basis set combined with different truncations of the excitation operator: only within RAS2 (0), with holes (h), with particles (p), and with hole and particle excitations (h, p). RAS-SF values are shown in Table 7.2 compared with SOCCs computed at the CASSCF level with the same active space and basis set [462, 464]. The results in Table 7.2 show that RAS-SF SOCCs systematically increase with the presence of hole and/or particle configurations. Concretely, hole configurations have a notably larger impact on the coupling constants, since relativistic effects are stronger in inner orbitals (closer to the nuclei than those conforming the RAS2 and RAS3 spaces). Qualitatively, RAS-SF couplings are in good agreement with CASSCF values reported by Furlani [464]. In particular, there is a good quantitative agreement with RAS-SF calculations including particle contributions ($\hat{R} = \hat{r}_0 + \hat{r}_p$), which can be related to the frozen core approximation used in the CASSCF calculations (see Ref. [464] for details).

Table 7.2: Singlet-triplet SOCC (in cm^{-1}) computed for different TM conformers with CASSCF and RAS-SF with different truncation of the excitation operator. All values obtained with an active space (RAS2) with 2 electrons in 2 orbitals and with the 3-21G basis set.

(α, β)	CASSCF[a]	RAS-SF(0)	RAS-SF(h)	RAS-SF(p)	RAS-SF(h,p)
(0, 0)	0.004	0.002	0.001	0.003	0.003
(0, 90)	0.023	0.000	0.049	0.020	0.068
(35, 35)	0.211	0.182	0.263	0.203	0.288
(145, 35)	0.177	0.166	0.225	0.193	0.253
(150, 60)	0.390	0.317	0.461	0.389	0.534
(45, 60)	0.690	0.547	0.827	0.673	0.959
(35, 90)	0.691	0.536	0.810	0.669	0.947
(90, 90)	1.760	1.334	2.020	1.747	2.439

[a]From Ref. [464].

As noted in a previous work [268], calculated SOCs can be rather sensitive to the treatment of static electron correlation. In order to explore the impact of the active space on the

singlet-triplet SOCC magnitude, we perform RAS-SF calculations with the cc-pVQZ basis set and with different RAS2 spaces: (2e, 2o), (4e, 3o), and (6e, 5o). In general, RAS-SF SOCCs increase with the size of the active space (Table 7.3), which can be related to both the inclusion of deeper orbitals and with the mixing of a larger variety of orbital contributions. Similar increments with the size of the correlated space are also obtained with other basis sets (Table 29).

Table 7.3: Singlet-triplet SOCC (in cm^{-1}) computed for different TM conformers at the RAS-SF/cc-pVQZ level with three different RAS2 spaces.

(α, β)	(2e, 2o)	(4e, 3o)	(6e, 5o)
(0, 0)	0.002	0.008	0.005
(0, 90)	0.074	0.069	0.058
(35, 35)	0.304	0.331	0.375
(145, 35)	0.398	0.419	0.515
(150, 60)	0.724	0.765	0.906
(45, 60)	1.130	1.201	1.341
(35, 90)	1.166	1.234	1.401
(90, 90)	2.982	3.055	3.536

Table 7.4: Systematic comparison of SOCCs (in cm^{-1}), effect of the RAS2 active space indicated in parenthesis computed with the 6-31G basis set.

(α, β)	(2e,2o)	(4e,3o)	(6e,5o)
(0,0)	0.0028	0.0097	0.0070
(90,0)	0.0733	0.0662	0.0557
(90,90)	2.6817	2.7177	3.0927
(150,60)	0.6029	0.6335	0.7300
(145,35)	0.3000	0.3165	0.3784
(35,90)	1.0396	1.0869	1.2053
(45,60)	1.0427	1.0909	1.1965
(35,35)	0.3025	0.3222	0.3562

Overall, these results demonstrate the good performance of our methodology, quantify the importance of hole and particle contributions, and the dependence on the RAS2 space and basis set. Moreover, our analysis unveils the relationship between computed couplings and biradical nature of the singlet state, which is complementary to the rationalization by Furland and King based on the ionic character.

7.5.2 Single bond dissociation: ethylene twist

Bond dissociation processes can be seen as the continuous transition from a closed-shell configuration to an open-shell electronic structure. This is the case of the stretching of a single (or multiple) bond or the torsion of a double (or multiple) bond. In the following, we explore the molecular twist of ethylene, representing the simplest model for π-bond dissociation. Molecular torsion from the planar (0° twist) ground state structure destroys the π-bonding interaction between the two sp^2 carbons, resulting in a single-electron occupation of two non-bonding degenerated orbitals at 90° twist (perfect diradical). The ethylene torsion model has been discussed in detail in numerous studies, highlighting the importance of

a balanced treatment of electron correlation [389, 390, 470, 471, 472] and its impact on the qualitative and quantitative estimation of singlet-triplet SOCs [275, 473, 399]. Here, in addition to test the performance of our methodology to the computation of SOCs along the twist mode of ethylene, we aim to establish connections between diradical character and the SOC strength. To this end, we employ singlet state geometries optimized at the CCSD(T)/cc-pVDZ level [399]. We perform RAS-SF calculations with the ROHF triplet state as the reference configuration, a RAS2 space containing 8 electrons in 7 orbitals (Figure 7.3), and the cc-pVTZ basis set.

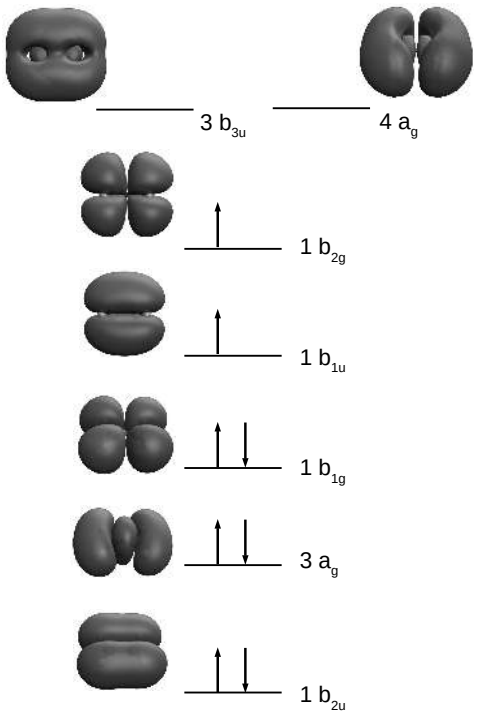


Figure 7.3: Triplet state orbital diagram corresponding to the RAS2 space (8e, 7o) of planar ethylene computed at the ROHF/cc-pVTZ level.

The twist distortion mode of C_2H_4 breaks the D_{2h} molecular point group symmetry of the ground state (planar, $\theta = 0°$) to D_2 ($0 < \theta < 90°$) and D_{2d} ($\theta = 90°$). The symmetry loss induced by the molecular twist distortion promotes a non-zero spin-orbit interaction of the ground state (S_0) with the lowest triplet (T_1). Finite SOC requires certain symmetry conditions, that is, the product of irreducible representations (Γ) of the two states (S_0 and T_1) and the SO Hamiltonian has to belong to the totally symmetric representation of the point group of the molecule. Since H^{SO} holds the symmetry of the molecular Hamiltonian and the triplet spin wave functions behave like the real rotations of the group, $\{R_k\}$ ($k = x, y, z$), the direct product of the irreducible representations of the singlet and triplet spatial wave functions must contain (at least) one of the $\Gamma(R_k)$. The two lowest electronic states at the D_{2h} geometry correspond to ^1A_g and $^3B_{1u}$, while rotations belong to B_{ng} ($n = 1, 2, 3$) representations. Hence, at $\theta = 0°$ SOC is symmetry forbidden. Similarly, at $\theta = 90°$ the symmetry of real rotations belong to E and A_2 representations, and SOC between $^1B_1/^3A_2$ states vanishes. On the other hand, for twisted D_2 structures the triplet spin wave functions hold B_n ($n = 1, 2, 3$) symmetries and the two states, ^1A and $^3B_1$, can couple through H^{SO}.

Accordingly, the RAS-SF singlet-triplet SOCC profile along the molecular twist presents values within the $0 \leq \text{SOCC} \leq 2.5$ cm^{-1} range, with exactly zero couplings at $\theta = 0°$ and 90° (Figure 7.4), which is an indication of the balanced treatment of electron correlation [399]. The total (mean-field) SOCC reaches a maximum of ~ 2.5 cm^{-1} at $50° < \theta < 60°$. The one-electron contribution follows a similar pattern, but with larger values, as the 2-electron component acts as a partial quencher of the 1-electron part. Detailed analysis of the relative importance of one and two-electron contributions to SOCs can be found elsewhere [457, 458].

To rationalize this behavior, we turn our attention to the electronic structure properties of S_0 and T_1. Regardless of the molecular torsion, the triplet state mainly corresponds to the singly occupancy of the two frontier orbitals (π and π^* at $\theta = 0°$), while the nature of the lowest singlet changes along the distortion. In general, the S_0 wave function can be approximated as the linear combination of two configurations, respectively featuring doubly occupancy of HOMO and LUMO,

$$\Psi(T_1) \sim \psi_{H,L} \tag{7.11}$$

$$\Psi(S_0) \sim c_{H^2}\psi_{H^2} + c_{L^2}\psi_{L^2} \tag{7.12}$$

where H^2 (L^2) subindex indicates the two-electron occupation of the HOMO (LUMO). It is worth noticing that the two contributions in $\Psi(S_0)$ present coefficients with opposite signs ($c_{H^2}c_{L^2} < 0$). From eqs 7.11 and 7.12, and considering only the 1-electron part of H^{SO}, the singlet-triplet SOC can be written as the spin-orbit interaction between HOMO (ϕ_H) and LUMO (ϕ_L) weighted by the expansion coefficients:

$$\langle \Psi(S_0)|H^{SO}|\Psi(T_1)\rangle \sim (c_{H^2} + c_{L^2})\langle \phi_H|H^{SO}|\bar{\phi}_L\rangle \tag{7.13}$$

where the over-bar indicates β spin-orbital. At small torsion angles the S_0 wave function is dominated by the closed-shell configuration ψ_{H^2} ($|c_{H^2}| > |c_{L^2}|$), as indicated by the large and rather constant $c_{H^2} + c_{L^2}$ value in Figure 7.4 ($c_{H^2} > 0$). Therefore, SOC at nearly planar structures linearly depends on the magnitude of the orbital interaction (eq 7.13), which increases with the orbital torque. The enhancement of the orbital SOC with θ is confirmed by the trend in the 1-electron couplings computed at the CIS level ($c_{H^2} = 1$, $c_{L^2} = 0$). On the other hand, for $\theta > 40°$ the weight of the doubly excited configuration becomes important and increases rapidly with θ. As a result, the sum over the two coefficients in eq 7.13 decreases since $c_{H^2}c_{L^2} < 0$, and becomes zero at $\theta = 90°$. The increase of $|c_{L^2}|$ can be related with a larger diradical character, as the sum of coefficients nicely correlates with the number of unpaired electrons quantified by the Head-Gordon index (see I_C and I_U profiles in Figure 7.4). Therefore, the singlet-triplet SOCC along the molecular twist of ethylene can be rationalized as the interplay between two factors, that is, the spin-orbit interaction between the frontier orbitals (promoting SOC) and the diradical character of the singlet (diminishing SOC), both increasing with θ.

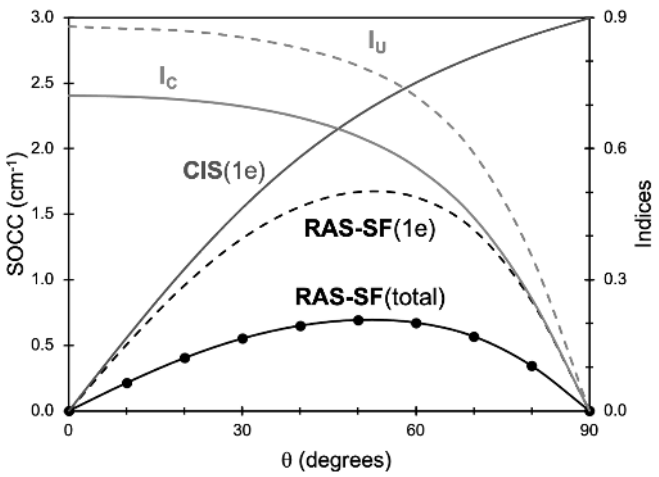


Figure 7.4: Singlet-triplet total and 1-electron RAS-SF/cc-pVTZ (black) and 1-electron CIS/STO-3G (blue) SOCCs, and $I_C = c_H + c_L$ and $I_U = (2 - N_U)/2$ indices (red) along the twist modes of ethylene. N_U correspond to the Head-Gordon index for the singlet state.

7.5.3 Double bond dissociation: ethylene stretch

Stretching of the C = C bond in ethylene involves the transition from a closed-shell structure (at the ground state equilibrium) to a tetraradical system formed by two triplet carbenes (at the dissociation limit) [474]. Spin-orbit interactions at large $CH_2 \cdots CH_2$ separations should be directly linked to the SOCs in the CH_2 fragment, while changes in the computed couplings upon association emerge from electronic interactions eventually resulting in the formation of σ and π bonds. Here, we investigate how electronic interactions inducing electron pairing (bond formation) between open-shell fragments, i.e., CH_2, impact interstate SOCs. Calculations along the C_2H_4 dissociation were done for two relative orientations of the CH_2 fragments: planar (D_{2h}, $\theta = 0°$) and twisted (D_2, $\theta = 50°$), with fixed CH_2 geometries to the ones employed in the previous section (different for $\theta = 0°$ and $\theta = 50°$). Electronic states and SOCs were obtained at the RAS-SF/cc-pVQZ level using the high-spin ROHF quintet state as the reference configuration with a (4e, 4o) RAS2 space formed by the singly occupied orbitals (σ, π, π^*, and σ^*) in ϕ_0, that is, $3a_g^1 3b_{3u}^1 1b_{2g}^1 1b_{1u}^1$ for the planar structure and $3a^1 3b_3^1 2b_2^1 2b_1^1$ for the twisted arrangement. This selection aims to generate a balanced description of the dissociation of the two bonds. RAS-SF calculations of the CH_2 unit have been done with the same basis set, a (2e, 2o) RAS2 space, and employing the ROHF triplet state as the reference configuration.

The CH_2 biradical holds a triplet ground state ($^3B_2$) with a close-lying excited singlet ($^1A_1$) computed at 0.63 (0.70) eV for the CH_2 fragment geometry from ethylene optimized at the 0° (50°) twist configuration, and with a rather large triplet-singlet SOCC of 11.15 cm^{-1} (11.02 cm^{-1}). Eigenstates of the (non-relativistic) Hamiltonian for two non-interacting CH_2 moieties, i.e., at large distances, are obtained as the spin-coupling of monomeric states. Hence, the three degenerated lowest states of $CH_2 \cdots CH_2$ at dissociation correspond to

$|T\rangle \otimes |T\rangle = \left|^{1,3,5}TT\right\rangle$, which can be written as ($M_S = 0$ states):

$$\left|^1TT\right\rangle = \frac{1}{\sqrt{3}}\{|T_+T_-\rangle + |T_-T_+\rangle - |T_0T_0\rangle\} \tag{7.14}$$

$$\left|^3TT\right\rangle = \frac{1}{\sqrt{2}}\{|T_+T_-\rangle - |T_-T_+\rangle\} \tag{7.15}$$

$$\left|^5TT\right\rangle = \frac{1}{\sqrt{6}}\{|T_+T_-\rangle + |T_-T_+\rangle + 2\,|T_0T_0\rangle\} \tag{7.16}$$

where fragment triplet state (T) subindices $+$, $-$, and 0 stand for the (local) M_S microstate (1, -1, and 0), $\left|^1TT\right\rangle = {}^1A_g$ (^1A), $\left|^3TT\right\rangle = {}^3B_{1u}$ ($^3B_1$), and $\left|^5TT\right\rangle = {}^5A_g$ (^5A) for the D_{2h} (D_2) symmetry. Above the TT manifold lie two triplets obtained as the spin coupling of local triplet and singlet states that correspond to $\{^3B_{2g}, {}^3B_{3u}\}$ and $\{^3B_2, {}^3B_3\}$ for the D_{2h} and D_2 symmetries, respectively.

Figure 7.5 tracks the evolution of these five states along the C=C stretching mode (higher energy excited states can be found in Appendix D). As the two CH_2 fragments approach each other, electronic interaction breaks the degeneracy between TT states, with the spin-singlet being the lowest state and converging to the ethylene closed-shell minimum at short distances. The triplet state also has a bonding profile, but presents a smaller binding energy, whereas the quintet is a dissociative state. The two triplet ST states remain degenerate and constant for CC distances larger than 3.5 Å. At shorter distances the two triplets are destabilized, with B_{2g} (B_3) lying at lower energies.

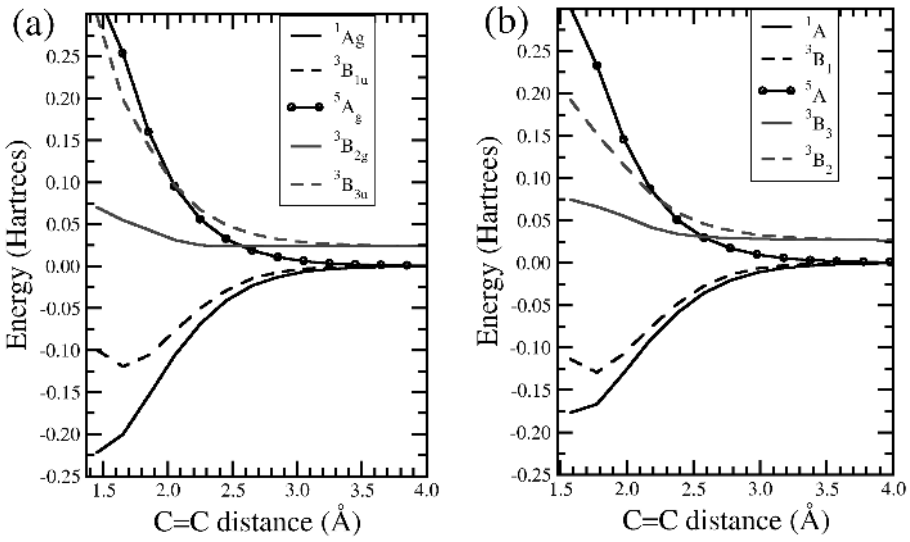


Figure 7.5: Dissociation curves for ethylene at 0° (a) and 50° (b) twist configurations computed at the RAS-SF/cc-pVQZ level. The graphs show the five lowest energy states selected for analysis.

Computed values of the Yamaguchi (N_Y) and Head-Gordon (N_U) indexes for the ground singlet state allow to identify and classify three regions along the dissociation curves (Figure 7.6). At short CC distance (< 2.5 Å) indexes display values lower than 2.0, suggesting a dominant closed-shell character. At larger intermediate distances N_Y and N_U rapidly increase, with values within the 2.0 – 3.0 range, indicating the diradical nature of the singlet ground state. Finally, towards dissociation the unpaired electron indices go from 3.0 to 4.0 (tetraradical region), corresponding to the ^1TT state, i.e., two CH_2 triplet fragments coupled

as an overall singlet. This classification is applicable to both analyzed symmetries. It is worth noticing that the energy curves of the excited triplet states ${}^3B_{2g}$ (${}^3B_3$) and ${}^3B_{3u}$ (${}^3B_2$) cross with that of the quintet state at the intermediate region (region D), with the lowest singlet state holding a diradical character.

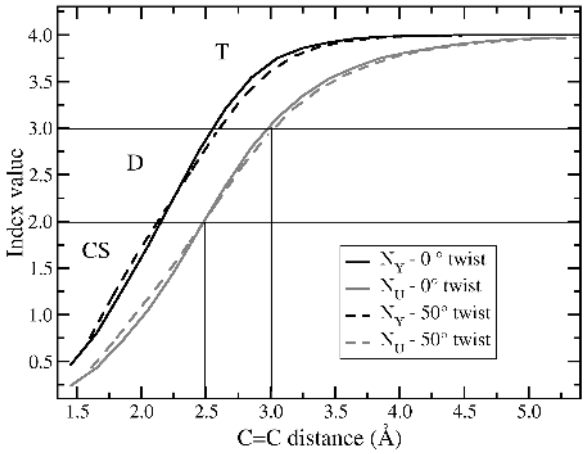


Figure 7.6: Yamaguchi (N_Y) and Head-Gordon (N_U) indexes for lowest singlet state along the dissociation of ethylene. The three discussed regions are indicated as CS (closed-shell), D (diradical), T (tetraradical).

In arrangements of non-interacting equivalent fragments, e.g., spatially isolated molecules, Frenkel excitons (linear combinations of local states) are degenerated to the states of individual fragments, but SOCs between pairs of degenerated states might not be invariant. Concretely, SOCCs (as defined in eq 7.10) between local states might be different than those between symmetry adapted eigenstates of the entire system Hamiltonian. To account for this additional degeneracy, it is convenient to define a SOCC considering not only the set of spin projections within each electronic configuration, but also the entire list of electronic states within the same manifold,

$$SOCC^2 = \frac{1}{N_I N_J} \sum_{I\in\Omega_I} \sum_{J\in\Omega_J} \sum_{M_S,M'_S} |\left\langle IM_S\right| H^{SO} \left|JM'_S\right\rangle|^2 \tag{7.17}$$

where I (J) runs over electronic states within the Ω_I (Ω_J) space with N_I (N_J) dimension, and M_S and M'_S indicate the different spin projections. The normalization factor $1/N_I N_J$ is chosen for convenience in order to enforce that SOCC of multiple non-interacting fragments is independent of the number of moieties. As the molecular units approach to each other, electronic interaction switches on, perturbing the nature of states and their properties, such as SOCCs in eq 7.17. Hence, the new SOCC definition directly informs about the overall change in spin-orbit interactions upon electronic coupling between fragments.

We apply eq 7.17 to the association of two methylene diradicals, up to the double bond formation (ethylene). In this case I and J run over TT and TS state manifolds, respectively. That is, $\Omega_I = \{{}^1TT, {}^3TT, {}^5TT\}$ ($N_I = 3$) and $\Omega_J = \{{}^3ST \pm {}^3TS\}$ ($N_J = 2$). SOCs for two methylene moieties infinitely separated obviously correspond to those interactions between local (CH_2) states. Therefore, the SOCC between TT and ST manifolds converge to the T/S (${}^3B_2/{}^1A_1$) SOCC value computed for CH_2 (Figure 7.7a).

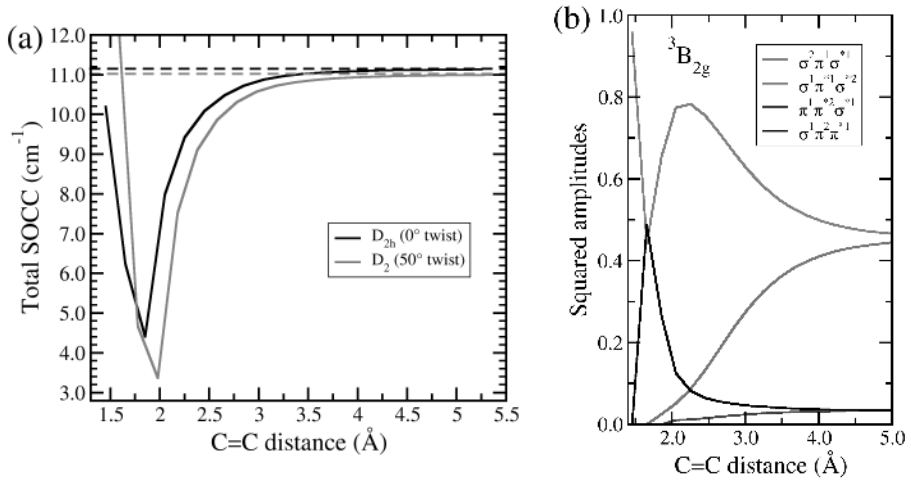


Figure 7.7: (a) SOCC profiles (in cm^{-1}, eq 7.17) between TT and ST manifolds of ethylene at 0° (black solid line) and 50° (red solid line) twist configurations along the C=C dissociation (RAS-SF/cc-pVQZ). Dashed horitzontal lines indicate the $^3B_2/^1A_1$ coupling in the CH_2 fragment. (b) Squared amplitudes of the main configurations for the $^3B_{2g}$ state along the stretching coordinate.

Approaching the two methylenes from dissociation (tetraradical region) to intermediate CC distances (diradicaloid region), there is a mild decrease of the computed SOCC. But as the ground state singlet weakens its open-shell character due to the consecutive formation of the σ and π bonds (distances smaller than 2.5 Å) the SOCC rapidly drops. The overall inter-manifold SOCC reaches a minimum at 1.6 – 2.0 Å distance and raises again for shorter separations. This behavior is general for all symmetry allowed SOCCs calculated between each pair of states from the two manifolds (Figures 7.8 and 7.9).

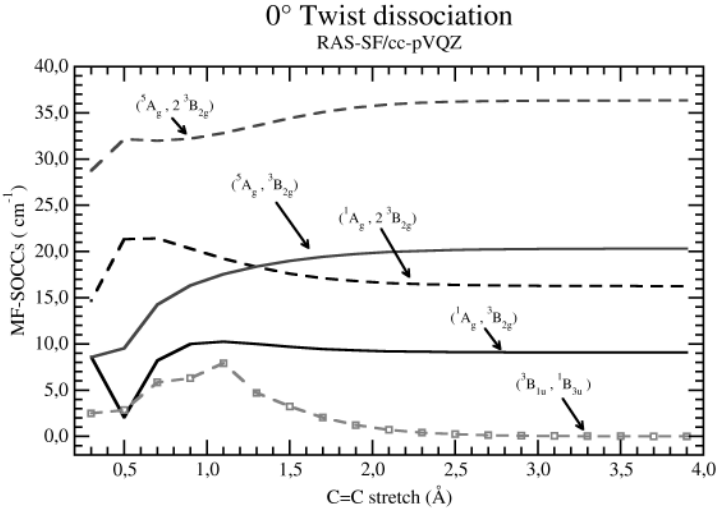


Figure 7.8: Plot of SOCCs values for ethylene at 0° twist configuration along the C=C stretching from equilibrium distance (RAS-SF/cc-pVQZ).

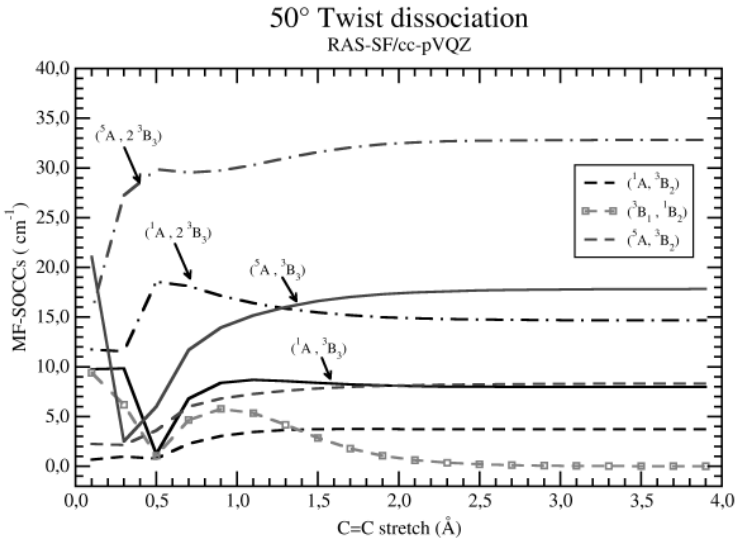


Figure 7.9: SOCC profiles for ethylene at 0° twist configuration along the C=C stretching from equilibrium distance (RAS-SF/cc-pVQZ).

In order to provide a qualitative rationalization for this behavior, we analyze the changes in the nature of electronic states with the inter-fragment distance. For the sake of simplicity, we focus on one of the ST states for the case of co-planar fragments, that is, the $^3B_{2g}$, which contributes to the total SOCC through spin-orbit interactions with the ^1A_g and ^5A_g states. Changes in the wave function character of $^3B_{1u}$ display an analogous and similar behaviour and are not showed here. At large separations, the $^3B_{2g}$ state is obtained as a nearly even contribution of two configurations: $\sigma^2\pi^1\sigma^{*1}$ and $\sigma^1\pi^{*1}\sigma^{*2}$ (Figure 7.7b), which can couple to the lowest quintet state ($\sigma^1\pi^1\pi^{*1}\sigma^{*1}$) through σ/π^* and σ^*/π spin-orbit interactions, respectively.

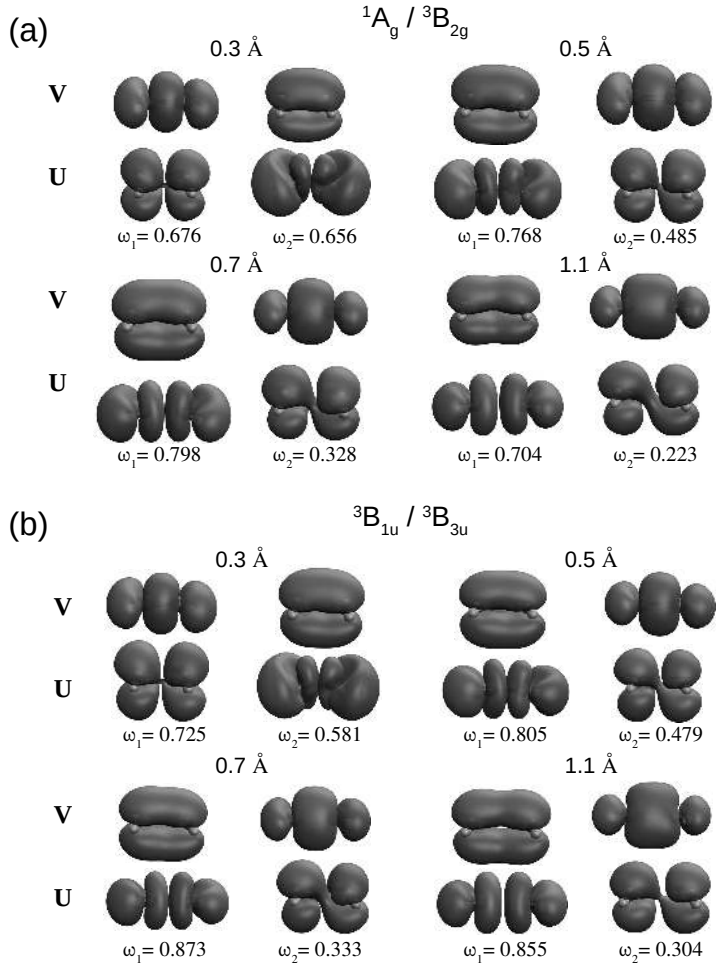


Figure 7.10: Two dominant NTO pairs of the spinless transition matrix (isovalue of 0.05) between $^1A_g/^3B_{2g}$ (a) and $^3B_{1u}/^3B_{3u}$ (b) states of planar ethylene computed for different CC distances at the RAS-SF/cc-pVQZ level. The two largest singular values are indicated as ω_1 and ω_2, respectively.

As the two units become closer to each other, the $\sigma^2\pi^1\sigma^{*1}$ ($\sigma^1\pi^{*1}\sigma^{*2}$) character of $^3B_{2g}$ increases (decreases). Simultaneously, there is an increase of the weight of the $\sigma^1\pi^2\pi^{*1}$, coinciding with the decay of the total SOCC and the suppression of the $^5A_g/^3B_{2g}$ SOCC (Figure 7.8). The role of this configuration in the $^3B_{2g}$ wave function decays again for C$\cdots$C distances shorter than 1.6 Å, concomitant with the SOCC raise. The change in character is also manifested in the main terms contributing to the interstate couplings as expressed by the spinless transition density matrix NTOs. For instance, whereas the σ/π^* coupling is the major contribution at shorter distances, the π/σ^* dominates at larger C$\cdots$C separations (Figure 7.10). Similar behavior is observed for the spinless transition density matrix NTOs obtained for the case with $50°$ torsion (Figure 7.11). Therefore, it becomes evident that the changes in the electronic structure character of states (related to σ and π-bond formation) have a large impact on interstate SOCs and the total SOCC between the two manifolds (TT and ST). Concretely, the mixing of the $\sigma^1\pi^2\pi^{*1}$ configuration into the triplet state wave function has a quenching effect on the computed SOCC.

The inherent collectivity of the transitions can be stated by the participation ratios (PRNTOs) (singular values and PRNTOs of the leading NTO pairs are displayed in Table 7.5). Pure

excitations acquire values close to 1, whereas transitions of mixed character display larger values [256]. At the beginning of the CS region (∼1.6 Å of separation) PRNTOS have large values close to 2, and decrease to values closer to 1 in going from CS to D regions. At larger C=C distances (T region) the participation ratios increase progressively, showing that excitations at dissociation are of mixed-character (PRNTO values larger than 2). This general behavior is observed in each of the components of the total SOCC.

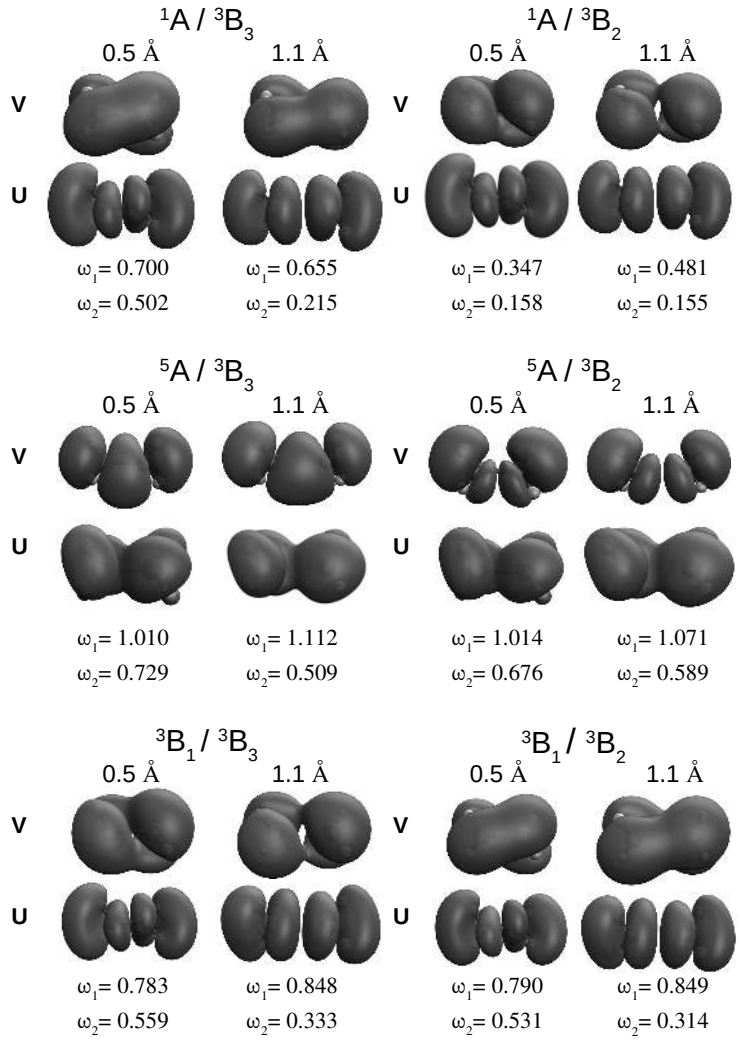


Figure 7.11: First dominant NTO pairs of the spinless transition matrix of ethylene (RAS-SF/cc-pVQZ) transitions at different stretch distances (isovalues of 0.05). The two largest singular values (ω_1 and ω_2) between the NTO pairs are presented below each pair.

Table 7.5: Singular values for the leading NTO pairs and participation ratios of the natural transition orbitals (PRNTOs) for the 0° and 50° twist configuration along dissociation, the equilibrium C=C distance for the planar (50° twist) geometry is 1.3516 (1.4794) Å. Here C=C (Å) regards to the stretch distance from equilibrium.

	0° planar			50° twist		
C=C (Å)	ω_1	ω_2	PRNTOs	ω_1	ω_2	PRNTOs
	coupling: $^1A_g/^3B_{2g}$			coupling: $^1A/^3B_3$		
0.3	0.6762	0.6560	2.0362	0.7649	0.4862	1.7298
0.5	0.7683	0.4855	1.7520	0.7000	0.5018	1.8816
1.1	0.7038	0.2232	1.3887	0.6547	0.2149	1.4203
3.9	0.3956	0.3871	2.3057	0.3942	0.3863	2.3280
	coupling: $^5A_g/^3B_{2g}$			coupling: $^5A/^3B_3$		
0.3	1.0782	0.6619	1.7201	1.1343	0.5529	1.5133
0.5	1.0479	0.6830	1.8598	1.0103	0.7290	2.0033
1.1	1.1302	0.4824	1.5947	1.1121	0.5094	1.6828
3.9	0.8820	0.8736	2.2986	0.8795	0.8716	2.3207
	coupling: $^3B_{1u}/^3B_{3u}$			coupling: $^3B_2/^3B_1$		
0.3	0.7251	0.5808	2.2122	0.9463	0.2305	1.1256
0.5	0.8046	0.4790	1.7693	0.7904	0.5306	1.8149
1.1	0.8552	0.3041	1.4813	0.8486	0.3136	1.5290
3.9	0.6807	0.6746	2.3066	0.6787	0.6731	2.3288

7.5.4 SOC in the ethylene-O_2 system

Molecular oxygen exhibits a spin-triplet ground state ($X^3\Sigma_g^-$), which makes it an efficient quencher of open-shell species [475, 476]. On the other hand, spin conservation in the reaction of O_2 with closed-shell molecules implies a triplet transition state.

Amongst organic compounds, olefins are particularly inclined to interact with oxygen, since carbon-carbon double bonds offer a specific point of attack. The reaction of molecular oxygen with the simplest olefin, i.e., ethylene, yields to either formaldehyde or to ethylene oxide (through the formation of a dioxetane-type addition complex) [477, 478]. The complexity of the oxidation reaction of ethylene is manifested by the large variety of products formed in slow combustion processes [479, 480, 481, 482]. Most of the studies take into consideration the more reactive singlet oxygen ($a^1\Delta_g$) [483] or $C_2H_4{}^+$ instead,[484] whereas the involvement of ISC has been rarely considered. SOC-mediated ISC between the ground state triplet and the excited singlet state in the isolated O_2 molecule is symmetry forbidden, but it can be activated by the interaction to other molecular species [485]. Perturbation of the oxygen cylindrical symmetry by molecular interactions modify the orbital properties and symmetry, promoting non-zero spin-orbit interactions with a magnitude depending on the geometry of the molecule-O_2 system [486].

Our purpose here is to explore possible mechanisms in which $C_2H_4 \cdots O_2$ interactions promote ISC through SOC. More generally, we aim to bring new results for the understanding of the ethylene oxidation reaction and shed light on the properties of the olefin-O_2 linkage. To this end, we explore a variety of possible geometrical arrangements for the $C_2H_4 \cdots O_2$ system (Figure 7.12). Molecular geometries of C_2H_4 and O_2 moieties in all studied dimers have been fixed to those optimized at the CCSD(T) level with the cc-pVDZ and 6-311+G(3df,2p) basis, respectively [399, 487]. Intermolecular distance in all dimers has been fixed at 2.0

Å, as indicated in Figure 7.12. SOCs have been computed at the RAS-SF level with the ROHF triplet configuration as the reference, (4e, 3o) in the RAS2 space, corresponding to the orbitals resulting from the interaction of the ethylene's HOMO and the two singly occupied π-orbitals in O_2, and the cc-pVQZ Dunning's basis set.

The ground state of all ethylene-oxygen model dimers present a triplet spin-multiplicity, with the two unpaired electrons mostly localized on O_2 (Figure 7.12). It is worth noticing that the **eclipsed** conformer also exhibits unpaired electrons on the ethylene unit. In most of the dimers, the computed triplet-singlet (TS) energy differences deviate considerably with respect to the gap obtained for the isolated O_2 molecule (1.262 eV), specially in those arrangements promoting the interaction between the π-orbitals of the two moieties (Table 7.6). The TS gap becomes considerably smaller in the **slip-stacked** and **in-plane-1** arrangements, and specially for the **eclipsed** disposition. On the other hand, intermolecular interactions are much weaker in the **inline** and **in-plane-2** forms, with energy gaps about 0.2 and 0.4 eV smaller than in molecular oxygen, respectively. The **crossed** conformer, with orthogonal π-systems, exhibits a TS gap very close to that in the O_2 molecule.

Most of the dimers show very strong singlet-triplet spin-orbit interactions, with SOCCs ranging from 76 to 103 cm^{-1} in all arrangements except those with C_{2v} symmetry (**eclipsed** and **crossed**), with vanishing couplings between the lowest triplet and singlet states. These results can be related to the SOC patterns observed in the interaction of O_2 with other molecular species [486].

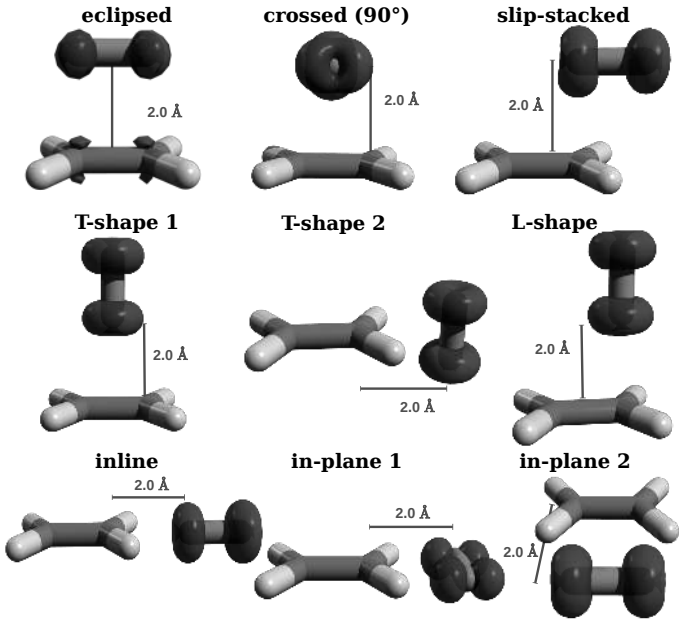


Figure 7.12: Ethylene-oxygen dimer structures with the representation of the FOD of the ground state triplet (blue isosurface).

Table 7.6: SOCC (in cm^{-1}) and vertical energy gaps (in eV) between the lowest triplet and singlet states of the ethylene-oxygen dimer structures in Figure 7.12 computed at the RAS-SF/cc-pVQZ level.

dimer	SOCC	TS gap
eclipsed	0.000	0.111
crossed	0.000	1.127
slip-stacked	99.301	0.167
T-shape-1	96.604	0.671
T-shape-2	88.293	0.593
L-shape	97.294	0.630
inline	97.086	0.830
in-plane-1	103.246	0.258
in-plane-2	76.316	1.079

Next, we explore the dependence of SOCCs upon symmetry loss along the relative coplanar torsion of the two molecules connecting the **eclipsed** and **crossed** structures in Figure 7.12. Therefore, we take CCOO torsion angles within the $0° \leq \theta \leq 90°$ range, with $\theta = 0°$ and $90°$ indicating parallel (**eclipsed**) and orthogonal (**crossed**) alignment between the C=C and O=O bonds, respectively. Whereas $0°$ and $90°$ conformers belong to the C_{2v} point group symmetry, intermediate angles reduce the symmetry of the system to C_2. Hence, for the sake of clarity, we employ orbital and state C_2-symmetry labels in all structures. Electronic states and couplings are obtained at the same level of computation as above, with the lowest ^3B ROHF state as the reference configuration.

At the **crossed** conformation, the two π-systems are orthogonal to each other, with fragment π-orbital interactions symmetry forbidden. As a consequence, the lowest state of the dimer (X^3B) can be described as the product of the two local electronic ground states: O_2 (triplet) and C_2H_4 (singlet), corresponding to the $9a^2 7b^1 10a^1$ configuration. About 1.1 eV above X^3B, lie two nearly degenerated singlet states respectively with A and B symmetries (Figure 7.13a). wave function analysis of these two states allow us to unequivocally identify them to the twofold excited singlet state of molecular oxygen ($a^1\Delta_g$). Therefore, we label them as $^1A(\Delta)$ and $^1B(\Delta)$ states, respectively. Finally, at higher energies, there is a singlet and triplet A-symmetric pair of excited states, mainly obtained as the electronic excitation from the ethylene's HOMO-like orbital ($9a$ in the dimer) to one of the ground state SOMOs ($7b$ of the dimer). In other words, these are $C_2H_4 \rightarrow O_2$ charge transfer (CT) states, and will be labelled as $^{1,3}A(\mathrm{CT})$.

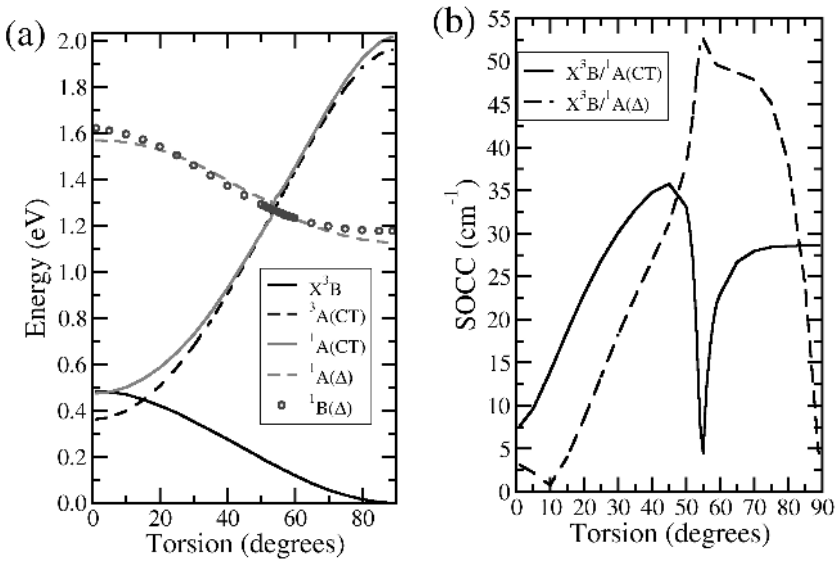


Figure 7.13: (a) Energy profiles (in eV) of low-lying states, and (b) SOCC (in cm^{-1}) between X^3B and the two lowest A-symmetry singlets, $^1A(\Delta)$ and $^1A(\mathrm{CT})$ states, along the coplanar torsion between **eclipsed** and **crossed** conformers computed at the RASCI/cc-pVQZ level of theory. Energies are given with respect to the ground state triplet at the **crossed** structure.

Reducing the torsion angle from the orthogonal disposition allows the $\pi - \pi$ orbital interactions (and orbital mixing) between the two fragments. As a consequence, there is a mild destabilization of the low-lying states localized on the O_2 moiety ($0.4 - 0.5$ eV between **crossed** and **eclipsed** geometries). The energy profiles of $^1A(\Delta)$ and $^1B(\Delta)$ are nearly parallel to X^3B, resulting in almost constant TS gaps. On the other hand, torsion-induced orbital overlap largely stabilizes CT states, which cross with $^1A(\Delta)$ and $^1B(\Delta)$ profiles at $\sim 55°$ and become nearly degenerate with the ^3B state close to the **eclipsed** conformation. In fact, the intersection between the two states with the same symmetry, $^1A(\Delta)$ and $^1A(\mathrm{CT})$, along our torsion model should be more properly identified as an avoided crossing, involving a very strong character mixing at the region with small energy gaps.

The $^3B/^1B$ couplings in the studied C_2-pathway are symmetry forbidden, hence, we focus our analysis on the couplings between X^3B and the two lowest A-symmetry excited singlets. Computed SOCs are greatly effected by state character mixing along the molecular torsion (Figure 7.13b). At the **crossed** structure the SOCC of the lowest triplet with $^1A(\Delta)$ vanishes, in accordance with the zero spin-orbit interaction between $X^3\Sigma_g^-$ and $a^1\Delta_g$ in the isolated O_2 [486]. On the other hand, the magnitude of the SOCC with $^1A(\mathrm{CT})$ is rather large, 28.6 cm^{-1}. Decreasing the intermolecular torsion angle from the orthogonal disposition induces a fast increase of the $^1A(\Delta)$ SOCC, with a maximum value of 52 cm^{-1} at $55°$, concomitant with the deep decrease of the $^1A(\mathrm{CT})$ coupling and the avoided crossing between the two A-symmetry singlets. The fast increase in SOCC can be related with the important $^1A(\Delta)/^1A(\mathrm{CT})$ character mixing, which is maximum at the avoiding crossing (Figure 7.14). For smaller torsion angles, both couplings decay up to values lower than 10 cm^{-1} at the **eclipsed** form.

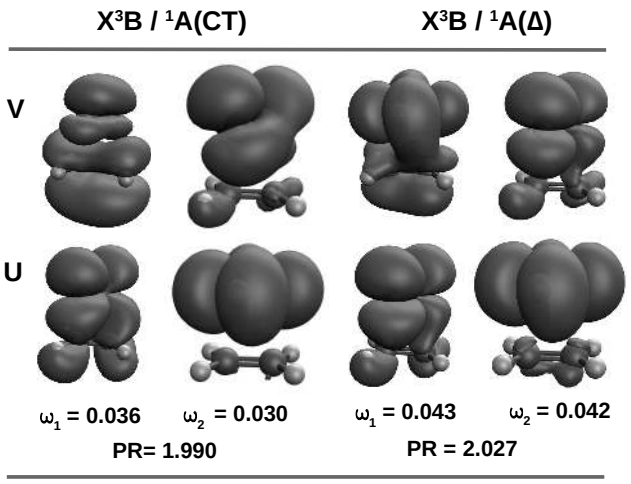


Figure 7.14: Dominant NTO pairs (isovalues of 0.05) of the spinless transition matrix between X^3B and the two lowest A-symmetry singlets for the $C_2H_4 \cdots O_2$ system with a 55° torsion angle. The participation ratios (PR) and the two largest singular values (ω_1 and ω_2) for the NTO pairs are also indicated.

Finally, we investigate the changes in SOCC upon increasing the intermolecular distance between C_2H_4 and O_2 fragments. For that, we consider two different regimes, that is, small (15°) and large (75°) torsion angles. For each torsion, we evaluate the SOCC between the ground state triplet and the lowest ^1A state for distances between 2 and 3 Å (Figure 7.15). For the larger torsion, separation of the two moieties systematically reduces the energy of the lowest triplet and singlet ($^1A(\Delta)$) states (Figure 7.15a). Simultaneously, the SOCC decays rather fast as the interaction between the two molecules decrease, recovering the triplet-singlet zero-coupling limit of O_2 at large distances (Figure 7.15b). At 15° of torsion, the lowest singlet is of CT character ($^1A(\mathrm{CT})$). As a consequence, separation of the two molecules induces a moderate increase of its energy, generating a large gap with the triplet ground state (Figure 7.15c). Along the dissociation, the energy profile of $^1A(\mathrm{CT})$ crosses with $^1A(\Delta)$ at 2.15 Å and with a higher ^1A state at 2.42 Å of separation. This higher A-symmetry state is mostly localized on the O_2 fragment and can be associated to the $b^1\Sigma_g^+$ state of molecular oxygen. Hence, we label it as $^1A(\Sigma)$. The state mixing around the two avoided crossings is clearly manifested in the SOCCs profiles (Figure 7.15d), with very strong coupling in the crossing with the $^1A(\Sigma)$ state.

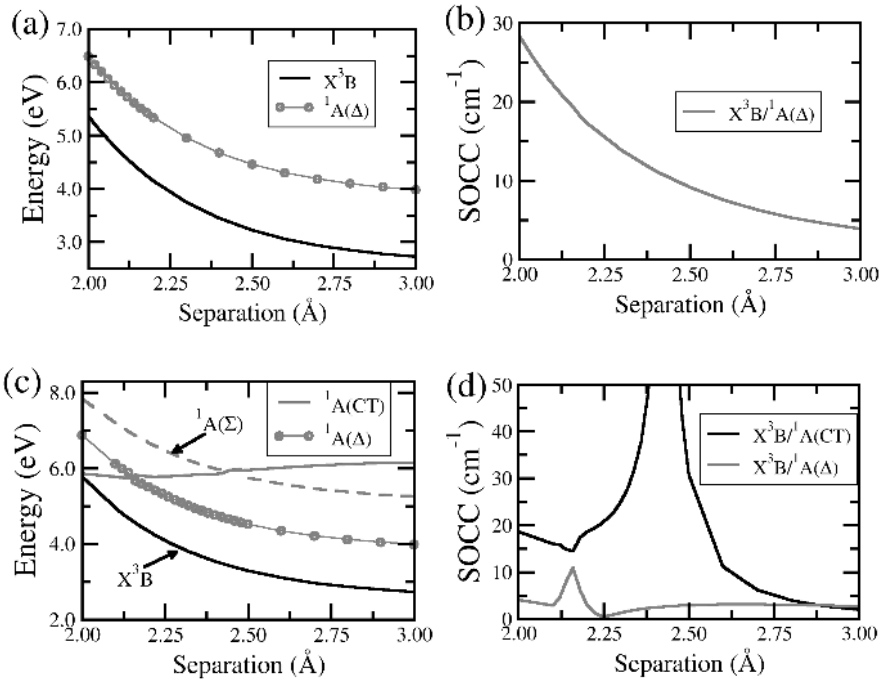


Figure 7.15: Energy (in eV) and SOCC (in cm^{-1}) profiles for the lowest lying X^3B and ^1A states of the $C_2H_4 \cdots O_2$ system at 75° (a and b) and 15° (c and d) torsion angles computed at different intermolecular separations.

Therefore, we can conclude that the interaction of O_2 with ethene can induce strong triplet-singlet spin-orbit interactions non-existing in the isolated oxygen molecule. The strength of the computed coupling is controlled by the orbital overlap between the two moieties and, as a consequence, it is very sensible to the relative disposition of the two molecules. Moreover, structural conformations close to the eclipsed disposition with short intermolecular distances facilitate the presence of low-energy CT (singlet) states able to strongly interact with the ground state triplet via SOC.

7.6 Conclusions

In this Chapter we have shown the capabilities of the RASCI method in conjunction with a recent implementation for the evaluation of the Breit-Pauli Hamiltonian in the evaluation of SOCs in open-shell systems. The present results demonstrate the suitability of the methodology for the computational analysis of challenging cases involving two or more unpaired electrons, the simultaneous dissociation of multiple bonds, or the interaction of molecules with ground states of different spin multiplicity. Our results indicate the importance of including hole and particle contributions in the RASCI wave functions for the evaluation of SOCs. In particular, hole configurations can have a large impact due to the involvement of inner orbitals. We have also shown how SOCs computed in diradical species are very sensible to structural parameters controlling the overlap between unpaired electrons, and the ionic character of the wave functions. Characterization and rationalization of SOCs in diradicaloid systems can be achieved by the representation of the dominant NTO pairs of the spinless transition density matrix. We have introduced a generalization of SOCC for the evaluation of couplings in aggregates of weakly interacting molecules, which allows to identify how intermolecular interactions tune SOCs with respect to those

in isolated species. We have computed this SOCC in the association of two methylene fragments to form ethylene. Moreover, we have been able to relate the obtained results with the character of the ground state singlet, going from a tetraradical (dissociation) to a diradical (intermediate distances) up to a closed shell system (at equilibrium), completing the formation of σ and π bonds. Finally, we have explored the role of intermolecular interactions in the triplet-singlet SOCs for the $C_2H_4 \cdots O_2$ dimer. Interaction between the π frontier orbitals of the two species results in large couplings, which are zero between the lowest singlet and triplet states of the isolated oxygen molecule. Moreover, for some arrangements with short intermolecular distances, e.g., eclipsed disposition, the sizeable π-orbital overlap triggers the presence of $C_2H_4 \rightarrow O_2$ CT states able to induce strong SOCs.

CHAPTER 8

General Conclusions

This thesis introduces new proposals in the realm of Quantum Chemistry computational methodologies, specifically focusing on the combination between WFT and DFT approaches, and on the use of an effective potential for treating the dynamic correlation missed in a CASSCF calculation. Through a comprehensive exploration of four distinct research studies, we have gained valuable insights into the capabilities and limitations of hybrid approaches. In fact, complete active space approaches have shown to be a WFT workhorse for method development. Due to their simplicity, they can be used for the construction of balanced mingled methods.

First, we investigated the potential of WFT-short-range DFT (*sr*DFT) for characterizing electronic states in atomic and molecular systems with strong open-shell character. Our findings revealed the challenges of achieving a balanced description between closed and open-shell configurations within the WFT-*sr*DFT framework. By recurring to alternative approaches that combine differently the exchange and correlation energy contributions, and by including and testing the incorporation of spin polarization effects into the DFT component, we identified that the combination of full WFT exchange energy with DFT correlation (WFT-*sr*DFTc) as the best strategy in terms of accuracy. Our preliminary results demonstrated the promise of this and other modifications, but further evaluations are essential for a broader range of molecular systems.

We have introduced a novel hybrid approach, WFT-*so*DFT, which integrates DFT correlation into wave function ansatzes by partitioning the natural orbital space based on NOONs. To this end, we employed the RASCI wave function approach with hole and particle truncation of the excitation operator for sampling the NO space. The method introduces a separation parameter ν, those NOs with occupations larger than ν are used to compute the WFT energy contributions, and the rest (small) occupations are used to estimate the DFT contribution through a density functional. Preliminary results suggest that low ν-values ($10^{-3} - 10^{-4}$) can yield reasonable correlation energies for correcting wave functions of the ground state in medium-sized molecules. However, systematic benchmarking is needed to understand the parameter's dependence and universality. Future efforts will involve comparing WFT-*so*DFT with other hybrid models and multiconfigurational perturbation theory. This method, requires a good description of the small occupation region, is suitable for use with large basis sets. RAS-*so*VWN results in atomic energies, bond dissociation, and torsion of ethylene show improvements over the pristine wave function (RASCI). However, the discrete distribution

of NOONs obtained with finite basis sets may introduce discontinuities in potential energy surfaces, particularly evident in N_2 dissociation, posing a limitation for characterizing large domains of potential energy surfaces.

Alternatively, we have explored the application of an effective potential to account for missing dynamic electron correlation in CASSCF wave functions. While the method showed potential for high accuracy in calculating correlation energies for small atomic systems, the challenge lies in establishing a robust and practical parametrization scheme for this potential. The impact of the effective potential on both the energy and electronic distribution of the system was identified, revealing discrepancies in the electronic distribution at short-range for negative potentials and improvements with positive potentials. In general, a suitable parametrization significantly reduces the errors in the CASSCF dissociation energy. Our findings underscore the significant impact of the effective potential on the system's energy and identifies the advantages of splitting the approach into intra- and inter-atomic contributions. In the case of H_2, parametrization results demonstrate that specific values of μ_{inter} consistently yield the lowest errors in dissociation energies, particularly when the intra-pair term is positive and maintained at values around 0.01. The splitting of the effective potential into inter- and intr-apair terms further elucidates their distinct effects on electronic distribution, with the intrapair term influencing short-range behavior. Furthermore, our findings suggest that utilizing a range of $\nu = 0.1$ to 0.50 would be optimal, with larger values potentially introducing excessive correlation energy.

These insights into the parametrization of the effective potential offer a foundation for future research and refinement, shedding light on its impact on both atomic and diatomic systems. While challenges persist, the ongoing developments in our laboratory aim to advance and improve this approach. This work contributes to the broader understanding of dynamic electron correlation treatment methods and provides valuable considerations for further exploration and enhancement.

Finally, the efficacy of the RASCI method, coupled with a recent implementation of the Breit-Pauli Hamiltonian, was demonstrated for calculating SOCs in open-shell systems. The methodology proved effective in handling challenging scenarios involving multiple unpaired electrons, simultaneous dissociation of multiple bonds, and interactions between molecules with different spin multiplicities. The characterization and rationalization of SOCs in diradicaloid systems were achieved through the representation of dominant NTO pairs in the spinless transition density matrix. The introduction of a generalized SOCC facilitated the evaluation of couplings in weakly interacting molecular aggregates, exemplified by the association of two methylene fragments to form ethylene. This approach provided insights into the transformation of the ground state singlet from a tetraradical during dissociation to a diradical at intermediate distances and eventually to a closed-shell system at equilibrium, marking the completion of σ and π bonds. The investigation also delved into the role of intermolecular interactions in triplet-singlet SOCs for the $C_2H_4 \cdots O_2$ dimer. The interaction between π frontier orbitals of the two species resulted in substantial couplings, particularly in arrangements with short intermolecular distances, such as the eclipsed arrangement, where significant π-orbital overlap induced $C_2H_4 \rightarrow O_2$ charge-transfer states capable of inducing strong SOCs.

We can conclude this thesis by recognizing that the introduced hybrid quantum chemistry methods remain a promising avenue for future research, but it is also evident that there are a considerable number of related issues to overcome. The identified issues encompass a spectrum of concerns, ranging from methodological intricacies to practical implementation challenges. For future perspectives intending to continue within this line of research, it is imperative for researchers and practitioners alike to delve deeper into these challenges, aiming to refine the hybrid methods to ensure their robustness and applicability.

Appendix A

Additional information for H_2

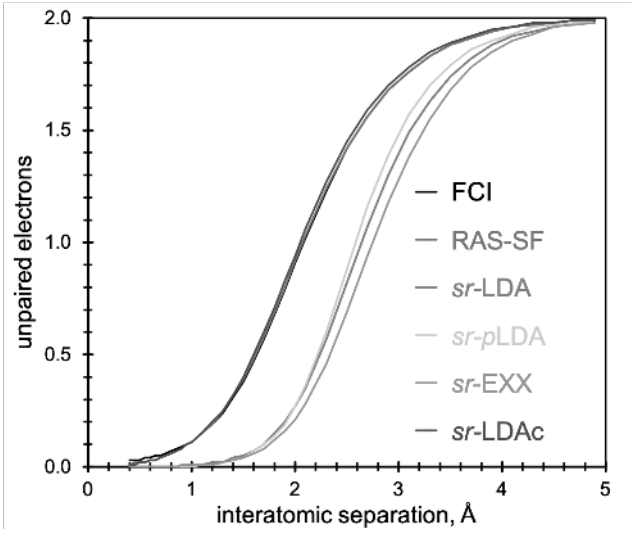


Figure 1: Number of unpaired electrons for the H_2 ground state along the molecular dissociation computed at the FCI, RAS-SF, RAS-*sr*LDA (*sr*-LDA), RAS-*sr*LDA with with spin polarized densities (*sr*-*p*LDA), with short-range HF exchange (*sr*-EXX), and with short-range correlation (*sr*-LDAc) levels and the 6-31G basis set.

Additional information for triatomic diradicals

Table 1: Adiabatic energy gaps (in eV) with respect to the $\widetilde{X}^3B_1$ ($M_s = 1$) ground state for CH_2 and NH_2^+ molecules computed with RAS(n, m)-SF-*sr*LDA, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively.

method	$\widetilde{a}^1A_1$	$\widetilde{b}^1B_1$	$\widetilde{c}^1A_1$
CH_2			
RAS(2,2)-*sr*LDA	-2.333	-1.057	-1.013
RAS(6,6)-*sr*LDA	-2.367	-1.068	-0.879
FCI[a]	0.483	1.542	2.674
Expt.[b]	0.390	1.425	
Expt. $-\Delta$ ZPE[c]	0.407	1.430	
NH_2^+			
RAS(2,2)-*sr*LDA	-1.952	-0.904	-0.861
RAS(6,6)-*sr*LDA	-2.171	-1.101	-1.071
CASSCF SOCI[a]	1.281	1.935	3.380
Expt.[d]	1.306		
Expt. $-\Delta$ ZPE[e]	1.286		

[a] TZ2P basis from Ref. [318]. [b] From Ref. [230]. [c] ZPE changes at the CISD level from Ref. [319]. [d] From Ref. [320]. [e] ZPE changes at the CISD level from Ref. [321].

Table 2: Adiabatic ST gaps (in eV) with respect to the $\widetilde{X}^1A_1$ ground state for SiH_2 and PH_2^+ molecules computed with RAS(n, m)-SF-*sr*LDA, where n and m indicate the number of electrons and orbitals in the RAS2 space, respectively.

method	$\widetilde{a}^3B_1$	$\widetilde{A}^1B_1$	$\widetilde{B}^1A_1$
SiH_2			
RAS(2,2)-*sr*LDA	2.940	2.175	3.162
RAS(6,6)-*sr*LDA	2.879	2.111	3.094
CASSCF SOCI[a]	0.871	1.992	3.486
Expt.	0.91[b]	1.928[c]	
Expt. $-\Delta$ ZPE[d]	0.90	1.92	
PH_2^+			
RAS(2,2)-*sr*LDA	3.218	2.256	3.124
RAS(6,6)-*sr*LDA	3.069	2.181	3.077
CASSCF SOCI[a]	0.760	2.009	3.686
Expt.[e]	0.75	1.92	
Expt. $-\Delta$ ZPE[f]	0.76	1.94	

[a] TZ2P basis from Ref. [318]. [b] From Ref. [325]. [c] From Ref. [326]. [d] ZPE changes at the CISD level from Ref. [323]. [e] From Ref. [327]. [f] ZPE changes at the CISD level from Ref. [324].

Appendix B

Additional information for polyene systems

We use the linear n-polyenes ($C_{2n}H_{2n+1}$, $n = 2-4$), to show the changes in NOON distribution affected by system size and level of theory. From considerations previously mentioned, in order to perform a correct orbital splitting and to incorporate DFT correlation, the distribution of occupations should behave similar to the momentum distribution for the HEG. Additionally, since the splitting of the NO space occurs at RAS3, it is fundamental that this subspace must be occupied (on the contrary, no DFT energy contribution is possible). Here we illustrate such characteristics through the use of CASCI and RASCI methodologies. To compute electronic wave functions through n spin-flip excitations from the ROHF single reference with $2n + 1$ spin multiplicity, i.e. single alpha-occupation for the $2n$ frontier pi-orbitals (RAS-nSF) the RAS2 space for each case is naturally defined as $2n$ electrons in the $2n$ pi-orbitals. Geometries have been optimized for the singlet ground state at the MP2/6-311(2+,2+)G(d,p) computational level.[125]

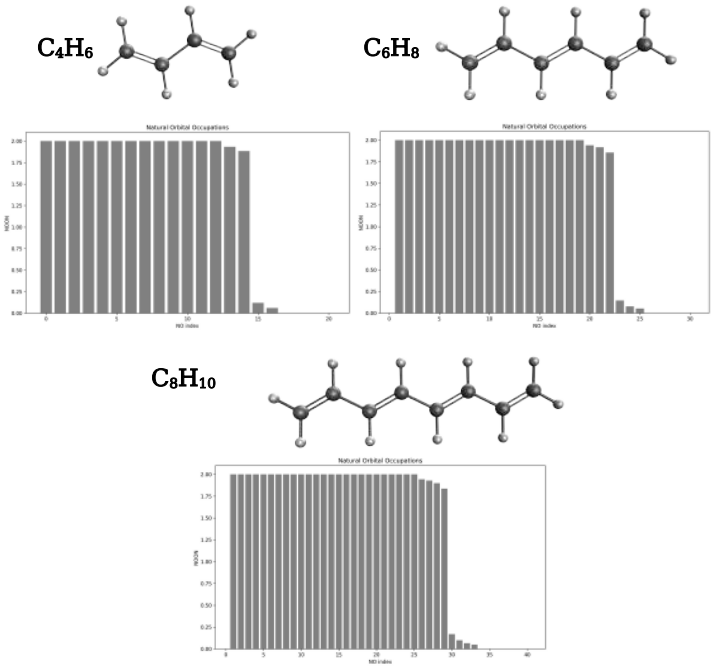


Figure 2: Illustrative NOONs computed at the RASCI(h,p)/cc-pVQZ level for $C_{2n}H_{2n+1}$ linear polyenes (n = 2-4).

Table 3: Sum of NO occupations for the ground singlet state ($1^1A_g^-$) in each RAS subspace for linear n-polyenes ($C_{2n}H_{2n+1}$).

basis	RAS1	RAS2	RAS3
2-polyene			
cc-pVDZ	25.9960	4.0002	0.0039
cc-pVTZ	26.0000	3.9970	0.0030
cc-pVQZ	26.0000	3.9965	0.0035
3-polyene			
cc-pVDZ	37.9930	6.0006	0.0064
cc-pVTZ	38.0000	5.9936	0.0064
cc-pVQZ	38.0000	5.9927	0.0073
4-polyene			
cc-pVDZ	49.9953	7.9963	0.0083
cc-pVTZ	50.0000	7.9902	0.0098
cc-pVQZ	50.0000	7.9889	0.0111

Table 4: Sum of NO occupations for the ground singlet state ($1^1A_g^-$) in each RAS subspace for linear n-polyenes ($C_{2n}H_{2n+1}$). Sums for the RASCI(p) and the RASCI(h,p) are the same, as for the RASCI(h) and CASCI calculations. Particle contributions are determinant for the presence of fractional occupancies at RAS3 space.

n	RAS1	RAS2	RAS3
CASCI			
2	26.0000	4.0000	0.0000
3	38.0000	6.0000	0.0000
4	50.0000	8.0000	0.0000
RASCI(h)			
2	26.0000	4.0000	0.0000
3	38.0000	6.0000	0.0000
4	50.0000	8.0000	0.0000
RASCI(p)			
2	26.0000	3.9965	0.0035
3	38.0000	5.9927	0.0073
4	50.0000	7.9889	0.0111

As the dimensions of the system increase, the NOONs profile of the singlet ground state assumes a likeness to the momentum distribution from HEG (see Fig.2). However, occupations at the overall distribution of occupations is sensible to the size of the basis set (See. Tables 3). While the smallest cc-pVDZ basis introduces some unbalanced distribution among the three orbital spaces, the use of larger basis sets seems to overcome this deficiency. In all cases, in going from the cc-pVTZ to the cc-pVQZ, occupations at RAS3 increase. Indeed, the introduction of particle contributions into the RASCI wave function is the driving force behind the emergence of non-zero occupancies in RAS3. This phenomenon is evident in Table 4. Notably, both RASCI(h) and CASCI methods fail to allocate any occupation to NO beyond the boundaries of RAS1 and RAS2.

Numerical results for atomic series

Complementary data for Atomic series

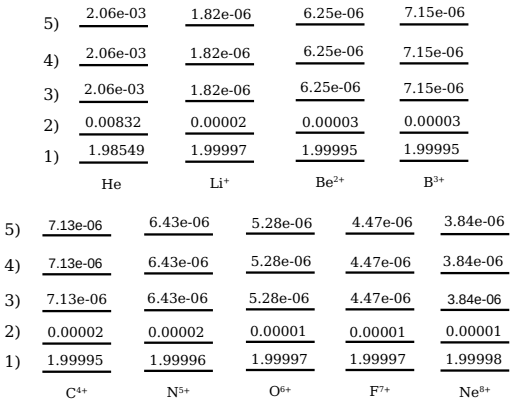


Figure 3: RASCI(h,p)/cc-pVDZ natural orbital occupation numbers (NOONs) representation for the isoelectronic He-series. Each horizontal line represents a natural orbital from the active space (RAS2(2e, 5o)), the numerical value of the respective NOON is placed above. Reported calculations where performed within this selection of the active space.

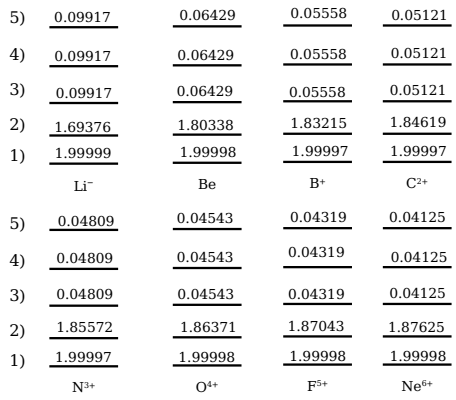


Figure 4: Full valence RASCI(h,p)/cc-pVDZ NOONs representation for the isoelectronic Be-series. Each horizontal line represents a NO from the active space (RAS2(4e, 5o)), the numerical value of the respective NOON is placed above.

In the following, a small change in notation is used: the labels 1S, 2S (2G). and 3S (3G) are used to designate, respectively, the use of a single NOON and cutting the NO space as in the original proposal by Savin; by using a single NOON and taking the sum of the rest of smaller occupations into the DFA with Savins's form of $n^+(r_s)$ (or Gori-Giorgi), or by using the sum of occupations criteria as described in the main text (again with the respective form of n^+, S or G).

Performance of different variants of the method

Exact correlation energies (in mhartrees) computed according to the definition of Wigner and Löwdin[415, 36] for the He- and Be-series and the respective percentage of errors obtained with the different versions of the method (and using the soVWN(S) functional form) are depicted in Table 5.

Table 5: [a]The 'exact' values of the correlation energies (E_c^{Exact}) were taken from Reference [3], and correspond to the difference between corrected experimental values and Hartree-Fock energies. A RAS2(2e, 5o) and RAS2(4e,5o) active spaces were used for the He- and Be-series respectively. An occupation threshold of 0.001 was chosen.

He-Series		%error (cc-pVTZ)			
Atom	$E_c^{Exact,a}$	RASCI	1S	2S	3S
He	-42.0	-30.64	0.42	-11.04	-11.04
Li^+	-43.5	-97.12	-48.70	-64.31	-64.31
Be^{2+}	-44.3	-96.55	-46.82	-63.63	-63.63
B^{3+}	-44.7	-94.82	-39.46	-59.46	-48.74
C^{4+}	-45.1	-93.72	-36.64	-58.32	-58.32
N^{5+}	-45.3	-92.92	-34.84	-58.06	-58.06
O^{6+}	-45.5	-92.47	-33.87	-58.71	-58.71
F^{7+}	-45.6	-92.11	-33.27	-46.99	-46.999
Ne^{8+}	-45.7	-91.80	-32.88	-60.01	-60.01
Be-Series		%error (cc-pVTZ)			
Atom	E_c^{Exact}	RASCI	1S	2S	3S
Li^-	-72.7	-70.13	-59.62	-63.57	-63.57
Be	-94.3	-60.19	-46.20	-49.65	-49.65
B^+	-111.4	-48.86	-3.91	-10.48	-10.48
C^{2+}	-126.4	-43.01	-18.95	-23.18	-23.18
N^{3+}	-140.5	-39.50	-21.43	-25.58	-25.58
O^{4+}	-154.1	-37.24	-20.92	-25.83	-25.83
F^{5+}	-167.3	-35.23	-20.12	-25.20	-25.20
Ne^{6+}	-180.4	-33.49	-19.38	-24.48	-24.48

In this case the two methods 2S and 3S give the same results. This is because the sum or occupations has a small value with essentially no differences in the two versions.

Comparison of correlation energies computed for the Be-series.

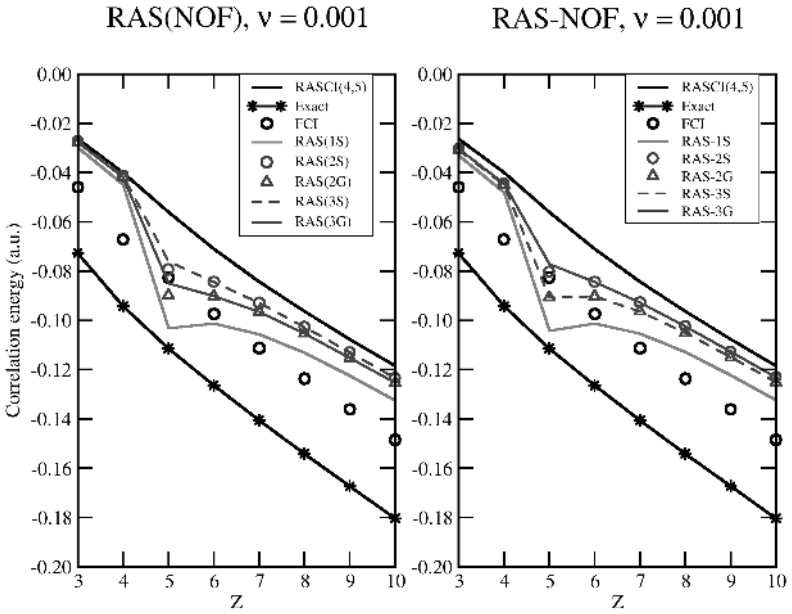


Figure 5: Correlation energies computed for the Be-Series with: *on-top* RAS(NOF) and RAS-NOF schemes. NOONs ($\nu_{threshold} = 0.001$) was chosen in order to get a comparison with the results reported with $\nu_{threshold} = 0.01$. Computed energies are close to the FCI energy than the previous case.

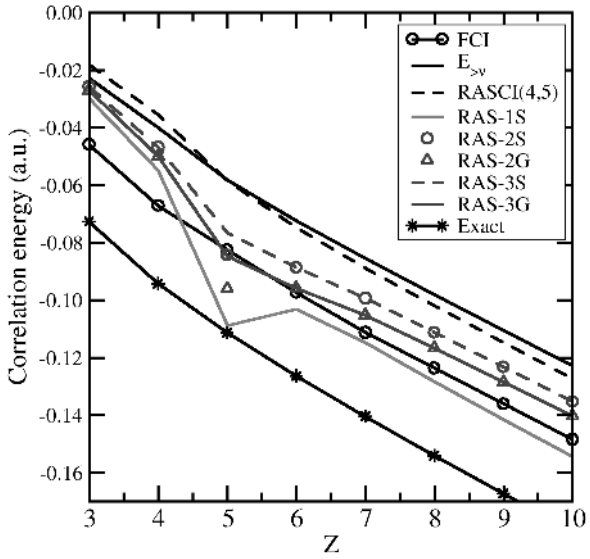


Figure 6: Correlation energies computed for the Be-Series within the RAS-NOF scheme $\nu_{threshold} = 0.01$. The use of a large basis set improves the calculation, however a similar overestimation as that obtained with the cc-pVDZ basis set remains.

Total energies computed with the cc-pVQZ basis set

Table 6: Total energies (in a.u.) for the He-series computed within different approaches as discussed in main text (with the soVWN(G) functional and $\nu = 0.001$), the cc-pVQZ basis set was used.

Atom	HF	RASCI(h,p)	CASSCF	FCI	$E > \nu$	soDFT	RAS-soDFT
He	-2.8615	-2.8864	-2.8974	-2.9024	-2.8931	-0.0140	-2.9072
Li^{+}	-7.2364	-7.2378	-7.2525	-7.2525	-7.2498	0.0000	-7.2498
Be^{+2}	-13.6113	-13.6143	-13.6302	-13.6304	-13.6264	0.0000	-13.6264
B^{+3}	-21.9862	-21.9904	-22.0054	-22.0054	-22.0017	0.0000	-22.0017
C^{+4}	-32.3611	-32.3661	-32.3811	-32.3812	-32.3736	0.0000	-32.3736
N^{+5}	-44.7360	-44.7417	-44.7565	-44.7566	-44.7492	0.0000	-44.7492
O^{+6}	-59.1108	-59.1169	-59.1312	-59.1313	-59.1242	0.0000	-59.1242
F^{+7}	-75.4857	-75.4921	-75.5060	-75.5061	-75.4993	0.0000	-75.4993
Ne^{+8}	-93.8605	-93.8671	-93.8809	-93.8810	-93.8743	0.0000	-93.8743

Table 7: Total energies (in a.u.) for the He-series computed within different approaches as discussed in main text (with the soVWN(S) functional and $\nu = 0.001$), the cc-pVQZ basis set was used.

Atom	HF	RASCI(h,p)	CASSCF	FCI	$E > \nu$	soDFT	RAS-soDFT
He	-2.8615	-2.8864	-2.8974	-2.9024	-2.8931	-0.0058	-2.8989
Li^{+}	-7.2364	-7.2378	-7.2525	-7.2525	-7.2498	0.0000	-7.2498
Be^{+2}	-13.6113	-13.6143	-13.6302	-13.6304	-13.6264	0.0000	-13.6264
B^{+3}	-21.9862	-21.9904	-22.0054	-22.0054	-22.0017	0.0000	-22.0017
C^{+4}	-32.3611	-32.3661	-32.3811	-32.3812	-32.3736	0.0000	-32.3736
N^{+5}	-44.7360	-44.7417	-44.7565	-44.7566	-44.7492	0.0000	-44.7492
O^{+6}	-59.1108	-59.1169	-59.1312	-59.1313	-59.1242	0.0000	-59.1242
F^{+7}	-75.4857	-75.4921	-75.5060	-75.5061	-75.4993	0.0000	-75.4993
Ne^{+8}	-93.8605	-93.8671	-93.8809	-93.8810	-93.8743	0.0000	-93.8743

Table 8: Total energies (in a.u.) for the Be-series computed within different approaches as discussed in main text (with the soVWN(G) functional and $\nu = 0.001$), the cc-pVQZ basis set was used.

Atom	HF	RASCI(h,p)	CASSCF	FCI	$E > \nu$	soDFT	RAS-soDFT
Li^{-}	-7.4248	-7.4429	-7.4515	-7.47073	-7.4477	-0.0684	-7.5160
Be	-14.5730	-14.6084	-14.6168	-14.64012	-14.6130	-0.0647	-14.6776
B^{+}	-24.2375	-24.2960	-24.2962	-24.31998	-24.2957	-0.0378	-24.3335
C^{+2}	-36.4083	-36.4830	-36.4807	-36.50565	-36.4806	-0.0235	-36.5041
N^{+3}	-51.0819	-51.1707	-51.1673	-51.19304	-51.1673	-0.0155	-51.1828
O^{+4}	-68.2569	-68.3589	-68.3550	-68.38053	-68.3550	-0.0165	-68.3715
F^{+5}	-87.9328	-88.0476	-88.0444	-88.0689	-88.0433	0.0000	-88.0433
Ne^{+6}	-110.1092	-110.2366	-110.2320	-110.25768	-110.2320	0.0000	-110.2320

Table 9: Total energies (in a.u.) for the Be-series computed within different approaches as discussed in main text (with the *so*VWN(S) functional and $\nu = 0.001$), the cc-pVQZ basis set was used.

Atom	HF	RASCI(h,p)	CASSCF	FCI	$E > \nu$	*so*DFT	RAS-*so*DFT
Li^-	-7.4248	-7.4429	-7.4515	-7.4707	-7.4477	-0.0460	-7.4937
Be	-14.5730	-14.6084	-14.6168	-14.6401	-14.6130	-0.0441	-14.6571
B^+	-24.2375	-24.2960	-24.2962	-24.3200	-24.2957	-0.0261	-24.3218
C^{+2}	-36.4083	-36.4830	-36.4807	-36.5057	-36.4806	-0.0164	-36.4970
N^{+3}	-51.0819	-51.1707	-51.1673	-51.1930	-51.1673	-0.0109	-51.1781
O^{+4}	-68.2569	-68.3589	-68.3550	-68.3805	-68.3550	-0.0117	-68.3666
F^{+5}	-87.9328	-88.0476	-88.0444	-88.0689	-88.0433	0.0000	-88.0433
Ne^{+6}	-110.1092	-110.2366	-110.2320	-110.2577	-110.2320	0.0000	-110.2320

H_2 bond dissociation

Complementary data for H_2 dissociation

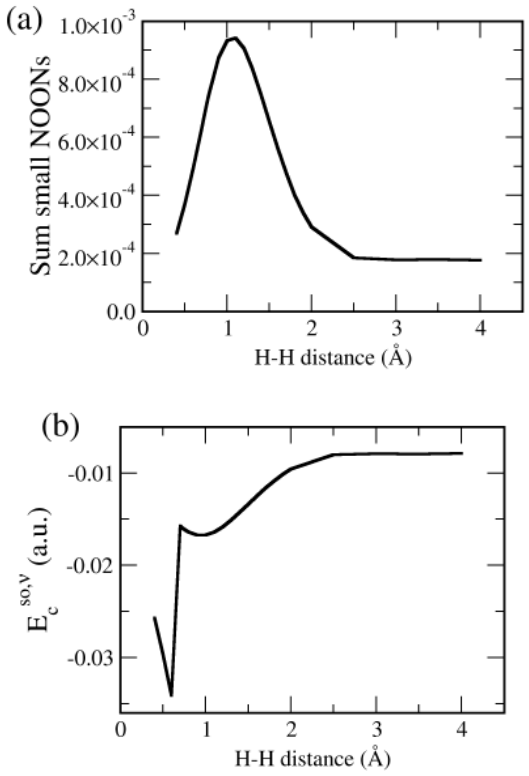


Figure 7: Complete data for (a) sum of small NOONs, and (b) *so*VWN energy (in a.u.) along the H_2 dissociation.

Basis set dependence

Table 10: Dissociation energies (in mHatree) for H_2 computed with different basis sets. Threshold for the sum of occupations used is 0.001. Experimental value is 164.5 mHartree.[4]

basis	RASCI	$\mathcal{E}_0^{\mathrm{WFA},\nu}$	soVWN(G)	FCI
cc-pVDZ	141.3	148.5	162.2	163.9
cc-pVTZ	150.5	159.3	160.2	171.1
cc-pVQZ	150.7	159.8	160.8	172.2

Potential energy surfaces for H_2

Potential energy surfaces for hydrogen within the 2S and 2G versions are depicted in Fig. 8. Results obtained by addition of the soDFT correlation energy to the pristine WFA energy behave similar to those in the RAS-soDFT scheme, improving the energy at bonding regions while converging to the WFT solution at long range.

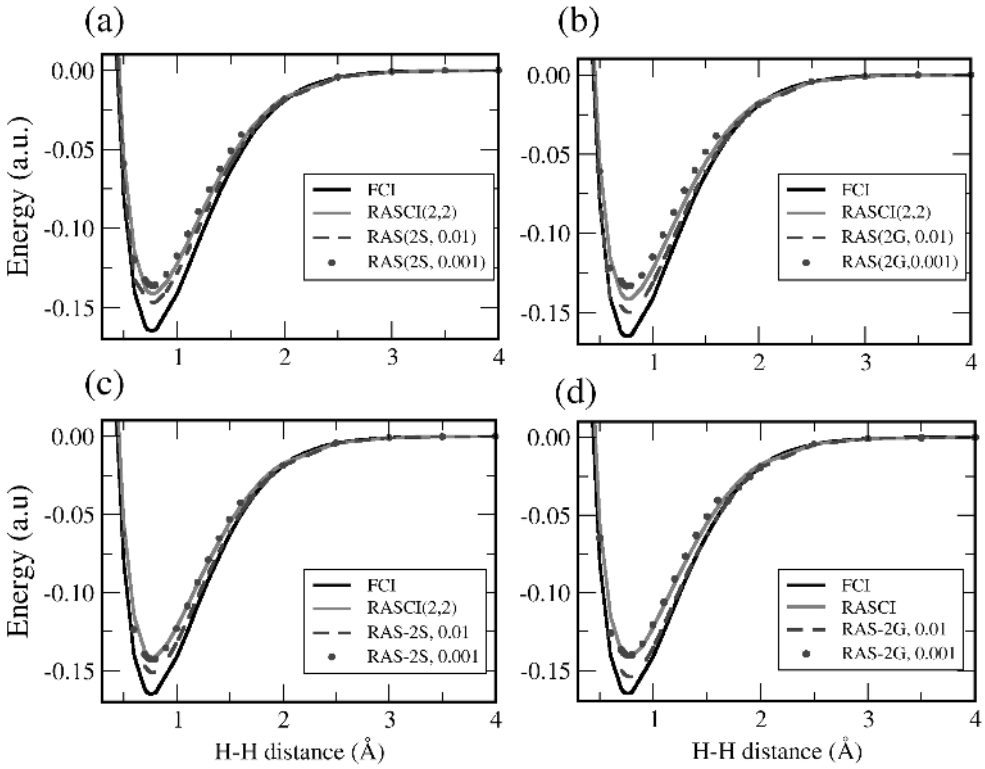


Figure 8: H_2 PES comparison among the $on-top$ (a and b) and the NO splitting (c and d) schemes, threshold values are indicated in the labels.

Table 11: Corresponding total energies for the H_2 PES computed in Figure 8 within the cc-pVDZ basis set in the *on-top* approach. The threshold value is indicated inside squared brackets below the RAS-*so*VWN functional labels.

Distance (Å)	RASCI (2e,2o)	*so*VWN(S) [0.01]	*so*VWN(S) [0.001]	*so*VWN(G) [0.01]	*so*VWN(G) [0.001]
0.4000	-0.9177	-0.9347	-0.9260	-0.9434	-0.9303
0.5000	-1.0535	-1.0731	-1.0626	-1.0831	-1.0672
0.6000	-1.1135	-1.1359	-1.1232	-1.1476	-1.1283
0.7000	-1.1360	-1.1463	-1.1360	-1.1517	-1.1360
0.7200	-1.1378	-1.1482	-1.1378	-1.1537	-1.1379
0.7400	-1.1390	-1.1495	-1.1390	-1.1550	-1.1390
0.7600	-1.1395	-1.1501	-1.1396	-1.1557	-1.1396
0.7800	-1.1396	-1.1503	-1.1397	-1.1559	-1.1397
0.8000	-1.1393	-1.1500	-1.1393	-1.1556	-1.1393
0.9000	-1.1325	-1.1435	-1.1326	-1.1492	-1.1327
1.0000	-1.1208	-1.1317	-1.1210	-1.1375	-1.1210
1.1000	-1.1069	-1.1176	-1.1071	-1.1233	-1.1073
1.2000	-1.0925	-1.1027	-1.0928	-1.1083	-1.0930
1.3000	-1.0784	-1.0882	-1.0790	-1.0935	-1.0792
1.4000	-1.0654	-1.0746	-1.0661	-1.0796	-1.0664
1.5000	-1.0536	-1.0622	-1.0545	-1.0669	-1.0550
1.6000	-1.0432	-1.0513	-1.0443	-1.0557	-1.0449
1.7000	-1.0343	-1.0418	-1.0418	-1.0459	-1.0459
1.8000	-1.0268	-1.0338	-1.0338	-1.0376	-1.0376
1.9000	-1.0206	-1.0272	-1.0272	-1.0307	-1.0307
2.0000	-1.0156	-1.0217	-1.0217	-1.0251	-1.0251
2.5000	-1.0026	-1.0078	-1.0078	-1.0106	-1.0106
3.0000	-0.9993	-1.0044	-1.0044	-1.0072	-1.0072
3.5000	-0.9985	-1.0036	-1.0036	-1.0064	-1.0064
4.0000	-0.9984	-1.0035	-1.0035	-1.0062	-1.0062

Table 12: Corresponding total energies for the H_2 PES computed in Figure 8 within the cc-pVDZ basis set in the RAS-*so*DFT approach. The threshold value is indicated inside squared brackets below the RAS-*so*VWN functional labels.

Distance (Å)	RASCI (2e,2o)	*so*VWN(S) [0.01]	*so*VWN(S) [0.001]	*so*VWN(G) [0.01]	*so*VWN(G) [0.001]
0.4000	-0.9177	-0.9498	-0.9296	-0.9498	-0.9338
0.5000	-1.0535	-1.0900	-1.0665	-1.0900	-1.0712
0.6000	-1.1135	-1.1548	-1.1275	-1.1548	-1.1325
0.7000	-1.1360	-1.1562	-1.1433	-1.1562	-1.1433
0.7200	-1.1378	-1.1582	-1.1451	-1.1582	-1.1452
0.7400	-1.1390	-1.1595	-1.1463	-1.1595	-1.1463
0.7600	-1.1395	-1.1603	-1.1468	-1.1603	-1.1469
0.7800	-1.1396	-1.1605	-1.1469	-1.1605	-1.1469
0.8000	-1.1393	-1.1602	-1.1465	-1.1602	-1.1465
0.9000	-1.1325	-1.1537	-1.1393	-1.1537	-1.1394
1.0000	-1.1208	-1.1417	-1.1270	-1.1417	-1.1271
1.1000	-1.1069	-1.1271	-1.1125	-1.1271	-1.1126
1.2000	-1.0925	-1.1117	-1.0974	-1.1117	-1.0976
1.3000	-1.0784	-1.0964	-1.0827	-1.0964	-1.0830
1.4000	-1.0654	-1.0821	-1.0692	-1.0821	-1.0695
1.5000	-1.0536	-1.0690	-1.0570	-1.0690	-1.0575
1.6000	-1.0432	-1.0574	-1.0463	-1.0574	-1.0469
1.7000	-1.0343	-1.0473	-1.0432	-1.0473	-1.0473
1.8000	-1.0268	-1.0387	-1.0349	-1.0387	-1.0387
1.9000	-1.0206	-1.0316	-1.0280	-1.0316	-1.0316
2.0000	-1.0156	-1.0258	-1.0224	-1.0258	-1.0258
2.5000	-1.0026	-1.0109	-1.0081	-1.0109	-1.0109
3.0000	-0.9993	-1.0074	-1.0046	-1.0074	-1.0074
3.5000	-0.9985	-1.0067	-1.0039	-1.0067	-1.0067
4.0000	-0.9984	-1.0065	-1.0037	-1.0065	-1.0065

Energy dependence on ν parameter

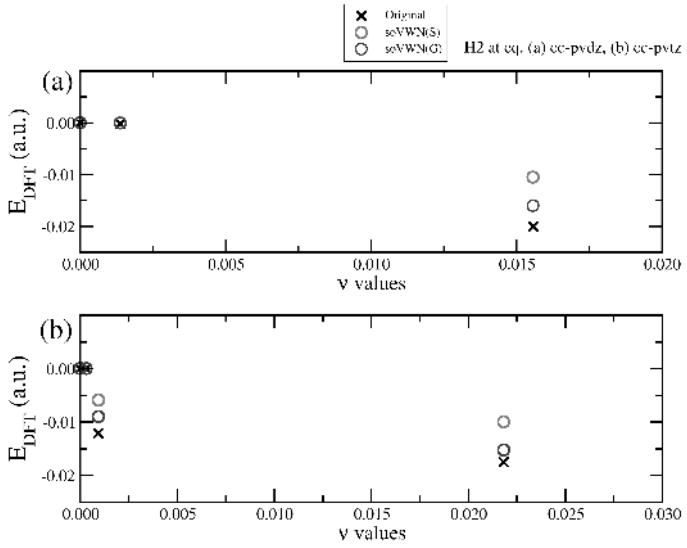


Figure 9: DFT energy dependence on the parameter nu for the H_2 ground state at $r_{HH} = 0.74$ Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets.

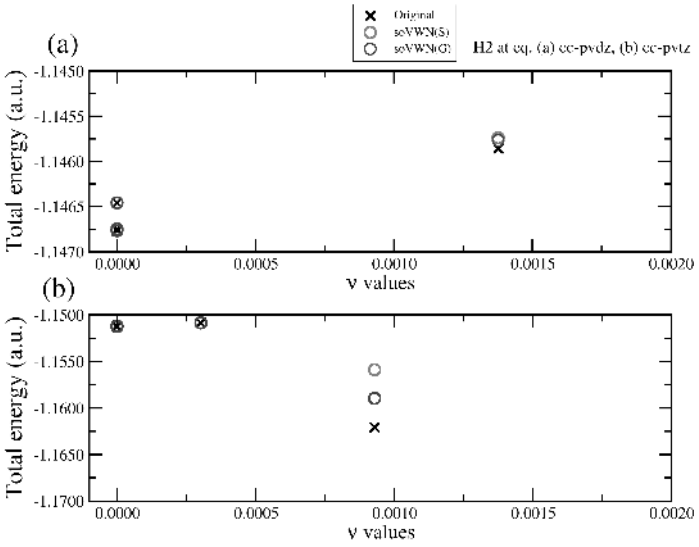


Figure 10: Total energy dependence on the parameter nu for the H_2 ground state at $r_{HH} = 0.74$ Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets.

C_2H_4 molecular torsion

Complementary data for C_2H_4 torsion

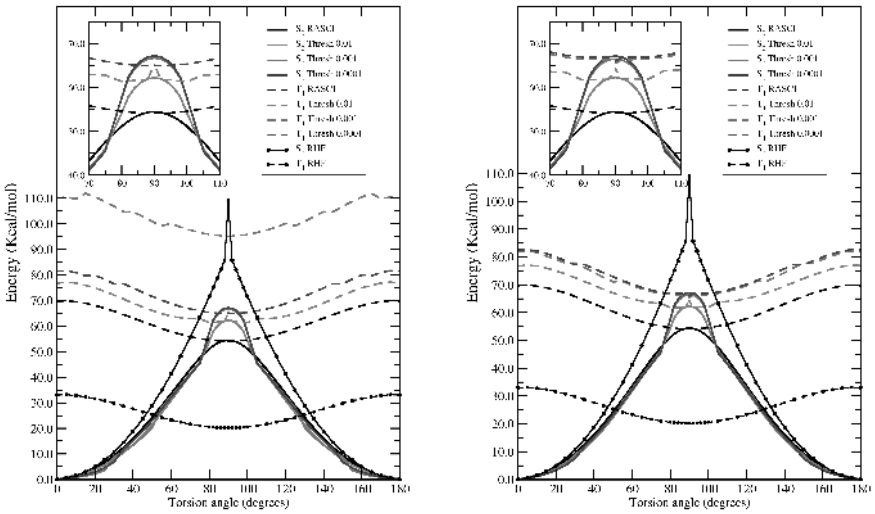


Figure 11: Ethylene potential energy surfaces for the singlet ground state and the lowest lying triplet state. Left: RAS-*so*DFT profiles obtained with the Savin (left) and Gori (right) $n_+(r_s)$ expressions, and with different values for the sum of occupations threshold.

Potential energy surfaces for C_2H_4

In the following tables, data for twisting modes of ethylene is summarized, values correspond to those for the PES shown above in Figure 11. Threshold values are indicated below method labels.

Table 13: PES data for the ethylene singlet state torsion (angle in degrees) corresponding values from Figure 11.

			RAS-*so*VWN(S)			RAS-*so*VWN(G)		
Angle	RHF	RASCI	0.01	0.001	0.0001	0.01	0.001	0.0001
0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	0.6004	0.6965	0.8785	0.6212	0.5585	1.0103	0.6338	0.5334
10	1.4736	1.4495	1.5123	1.1609	1.0793	1.5876	0.8722	0.9726
15	2.9272	2.6983	2.5540	2.0708	2.0080	2.5414	1.3240	1.7633
20	4.9585	4.4365	4.0286	3.5078	3.3509	3.8968	2.1712	2.9493
25	7.5639	6.6579	5.9362	5.3276	5.0640	5.6601	3.4952	4.5118
30	10.7385	9.3436	8.2831	8.8479	7.4987	7.8627	8.8165	6.8524
35	14.4762	12.4812	11.0881	11.5085	11.8411	10.5171	10.8936	11.8411
40	18.7696	16.0454	14.3511	14.6335	15.1732	13.6421	13.4664	15.1167
45	23.6098	19.9987	18.1225	18.2731	18.9822	17.3067	16.6290	18.8880
50	28.9866	24.2972	22.3895	22.3833	23.2178	21.4859	20.3313	23.0861
55	34.8878	28.8717	27.0143	28.1438	27.7108	26.0730	28.1877	27.5477
60	41.2995	33.6471	31.8398	32.7121	32.2916	30.9111	32.5489	32.0971
65	48.2056	38.4852	36.8160	37.3243	36.8850	35.9500	37.0293	36.6591
70	55.5877	43.2040	41.7733	41.7984	41.3842	41.0391	41.4282	41.1270
75	63.4249	47.5464	46.4796	46.0843	45.7078	45.9274	45.7956	45.4568
80	71.6932	51.1546	54.4929	57.6430	57.8124	54.1666	57.5301	57.6242
82	75.1152	52.2966	58.0321	62.0732	62.4372	57.7936	62.0481	62.2866
84	78.5999	53.2191	60.0464	64.2632	64.7527	59.8895	64.3323	64.6586
86	82.1454	53.8968	61.3265	65.5810	66.1583	61.2324	65.7316	66.1207
88	85.7464	54.3172	62.0419	66.2964	66.9553	61.9917	66.4846	66.9490
90	109.0452	54.4605	62.2866	66.5223	67.2251	62.2427	66.7105	67.2376
92	85.7464	54.3172	62.0419	66.2964	66.9553	61.9917	66.4846	66.9490
94	82.1454	53.8968	61.3265	65.5810	66.1583	61.2324	65.7316	66.1207
96	78.5999	53.2191	60.0464	64.2632	64.7527	59.8895	64.3323	64.6586
98	75.1152	52.2966	58.0321	62.0732	62.4372	57.7936	62.0481	62.2866
100	71.6932	51.1546	54.4929	57.6430	57.8124	54.1666	57.5301	57.6242
105	63.4249	47.5464	46.4796	46.0843	45.7078	45.9274	45.7956	45.4568
110	55.5877	43.2040	41.7733	41.7984	41.3842	41.0391	41.4282	41.1270
115	48.2056	38.4852	36.8160	37.3243	36.8850	35.9500	37.0293	36.6591
120	41.2995	33.6471	31.8398	32.7121	32.2916	30.9111	32.5489	32.0971
125	34.8878	28.8717	27.0143	28.1438	27.7108	26.0730	28.1877	27.5477
130	28.9866	24.2972	22.3895	22.3833	23.2178	21.4859	20.3313	23.0861
135	23.6098	19.9987	18.1225	18.2731	18.9822	17.3067	16.6290	18.8880
140	18.7696	16.0454	14.3511	14.6335	15.1732	13.6421	13.4664	15.1167
145	14.4762	12.4812	11.0881	11.5085	11.8411	10.5171	10.8936	11.8411
150	10.7385	9.3436	8.2831	8.8479	7.4987	7.8627	8.8165	6.8524
155	7.5639	6.6579	5.9362	5.3276	5.0640	5.6601	3.4952	4.5118
160	4.9585	4.4365	4.0286	3.5078	3.3509	3.8968	2.1712	2.9493
165	2.9272	2.6983	2.5540	2.0708	2.0080	2.5414	1.3240	1.7633
170	1.4736	1.4495	1.5123	1.1609	1.0793	1.5876	0.8722	0.9726
175	0.6004	0.6965	0.8785	0.6212	0.5585	1.0103	0.6338	0.5334
180	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 14: PES data for the ethylene triplet state torsion (angle in degrees) corresponding values from Figure 11.

			RAS-*so*VWN(S)			RAS-*so*VWN(G)		
Angle	RHF	RASCI	0.01	0.001	0.0001	0.01	0.001	0.0001
0	33.0100	69.6912	76.6754	81.9025	82.4736	76.6754	110.1781	81.5511
5	33.1438	69.9610	77.1711	82.1724	82.6932	77.3029	110.4605	81.7205
10	32.8258	69.5657	76.7256	81.5950	82.1033	76.8009	109.7828	80.9111
15	32.3094	68.9319	76.0039	80.3338	81.1997	75.9914	111.9414	79.7188
20	31.6138	68.0722	75.0438	79.1227	79.9447	74.9121	110.5170	79.9823
25	30.7645	67.0306	73.8830	77.7798	78.5893	73.6069	108.7850	78.3194
30	29.7915	65.8320	72.5777	77.8739	77.1962	72.1573	106.8397	76.6126
35	28.7285	64.5268	71.1784	76.2173	77.2715	70.6074	104.7564	76.6503
40	27.6105	63.1525	69.7539	74.3724	75.7780	69.0449	104.8505	74.9434
45	26.4732	61.7532	68.3546	72.6970	74.3097	67.5388	102.9053	73.2805
50	25.3514	60.3852	67.0431	71.0905	72.8727	66.1395	101.0165	71.6741
55	24.2779	59.0675	65.8697	71.2474	71.4984	64.9284	99.2594	70.1807
60	23.2830	57.8564	64.8719	69.7100	70.2120	63.9432	100.1066	68.8315
65	22.3938	56.7708	64.1126	68.5554	69.0637	63.2467	98.7261	67.6832
70	21.6337	55.8421	63.6106	67.5953	68.0848	62.8764	97.5463	66.7231
75	21.0222	55.1016	63.3910	66.9176	67.3380	62.8388	96.5800	65.9512
80	20.5744	54.5557	61.6779	66.5034	66.9239	61.3516	95.8709	65.3928
82	20.4437	54.3988	61.7218	66.4093	66.8674	61.4834	95.6638	65.2296
84	20.3415	54.2733	61.7908	66.3528	66.8549	61.6340	95.5007	65.1041
86	20.2683	54.1854	61.8473	66.3215	66.8800	61.7532	95.3689	65.0225
88	20.2212	54.1290	61.8787	66.2901	66.9051	61.8285	95.2873	64.9598
90	20.2119	54.1164	64.2381	65.5245	65.9952	65.2735	95.2622	64.9472
92	20.2212	54.1290	61.8787	66.2901	66.9051	61.8285	95.2873	64.9598
94	20.2683	54.1854	61.9289	66.3215	66.8800	61.7532	95.3689	65.0225
96	20.3415	54.2733	61.9289	66.3528	66.8549	61.6340	95.5007	65.1041
98	20.4437	54.3988	61.8975	66.4093	66.8674	61.4834	95.6638	65.2296
100	20.5744	54.5557	61.8662	66.5034	66.9239	61.3516	95.8709	65.3928
105	21.0222	55.1016	63.8616	66.9176	67.3380	62.8388	96.5800	65.9512
110	21.6337	55.8421	64.0122	67.5953	68.0848	62.8764	97.5463	66.7231
115	22.3938	56.7708	64.3825	68.5554	69.0637	63.2467	98.7261	67.6832
120	23.2830	57.8564	65.0162	69.7100	70.2120	63.9432	100.1066	68.8315
125	24.2779	59.0675	65.8822	71.2474	71.4984	64.9284	99.2594	70.1807
130	25.3514	60.3852	66.9615	71.0905	72.8727	66.1395	101.0165	71.6741
135	26.4732	61.7532	68.1789	72.6970	74.3097	67.5388	102.9053	73.2805
140	27.6105	63.1525	69.5092	74.3724	75.7780	69.0449	104.8505	74.9434
145	28.7285	64.5268	70.8835	76.2173	77.2715	70.6074	104.7564	76.6503
150	29.7915	65.8320	72.2640	77.8739	77.1962	72.1573	106.8397	76.6126
155	30.7645	67.0306	73.5629	77.7798	78.5893	73.6069	108.7850	78.3194
160	31.6138	68.0722	74.7426	79.1227	79.9447	74.9121	110.5170	79.9823
165	32.3094	68.9319	75.7467	80.3338	81.1997	75.9914	111.9414	79.7188
170	32.8258	69.5657	76.5248	81.5950	82.1033	76.8009	109.7828	80.9111
175	33.1438	69.9610	77.0519	82.1724	82.6932	77.3029	110.4605	81.7205
180	33.0100	69.6912	76.9640	81.9025	82.4736	76.6754	110.1781	81.5511

Energy dependence on ν parameter

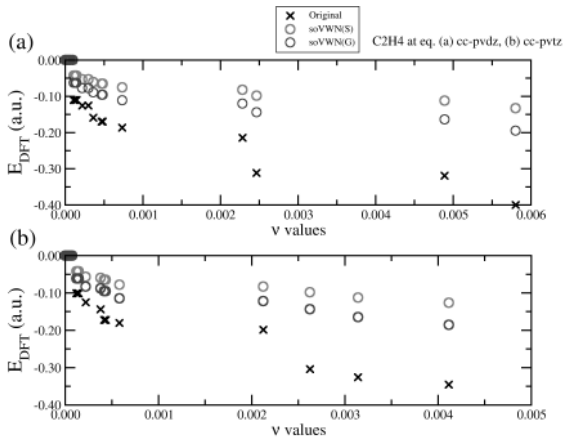


Figure 12: DFT energy dependence on the parameter nu for the C_2H_4 ground state (optimized at the CCSD(T)/cc-pVDZ level) computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets.

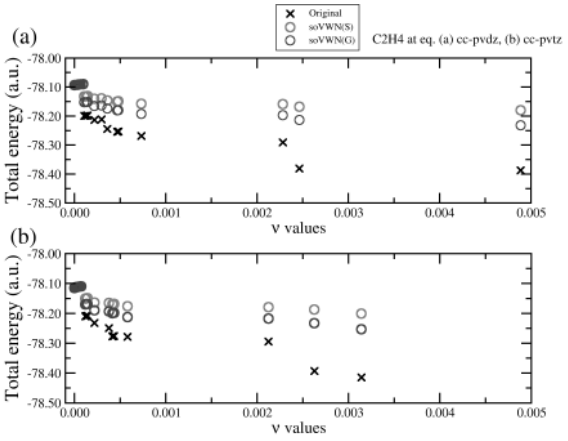


Figure 13: Total energy dependence on the parameter nu for the C_2H_4 ground state (optimized at the CCSD(T)/cc-pVDZ level) computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets.

N_2 triple bond dissociation

RASCI Natural orbital occupations computed for the N_2

In order to try to understand the overestimations made by the density functional in the region of dissociation, corresponding natural orbital occupations are plotted in Figures 15 and 16. From these it is observed that as the number of effective unpaired electrons increase, more natural orbitals acquire occupations that are not negligible and that are close to the chosen threshold. In the case of RAS(NOF0,2) the situation is not compensated by the sum of occupations. The problem arise due to the use of a single value for the occupation threshold. In fact this is the major problem with Savin's proposal, as the distribution of occupations from Figures 15 and 16 suggest, NO occupations can change along the PES. Furthermore, as the HG index shows (see Fig. 14), at dissociation it is expected a total amount of six unpaired electrons, during the transition, some occupations will achieve a value of 1.0 and the distribution profile will acquire a kind of "step" (see distribution of NO occupations at 2.5 Å of separation in Figure 15). At these long-range atomic separations, NO occupations in the RAS2 space tend to be close to the unity, and the density functional energy contribution converges to zero.

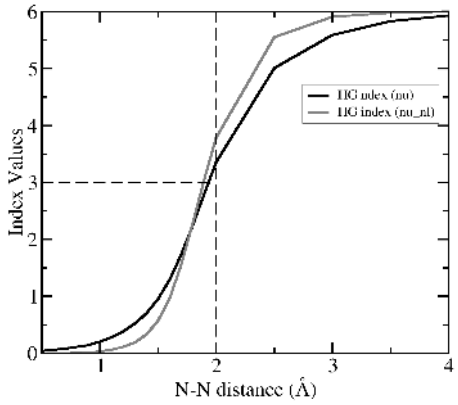


Figure 14: Computed values for the effective unpaired electrons accounted from the Head-Gordon "nu" and "nu_nl" indexes along the dissociation of N_2.

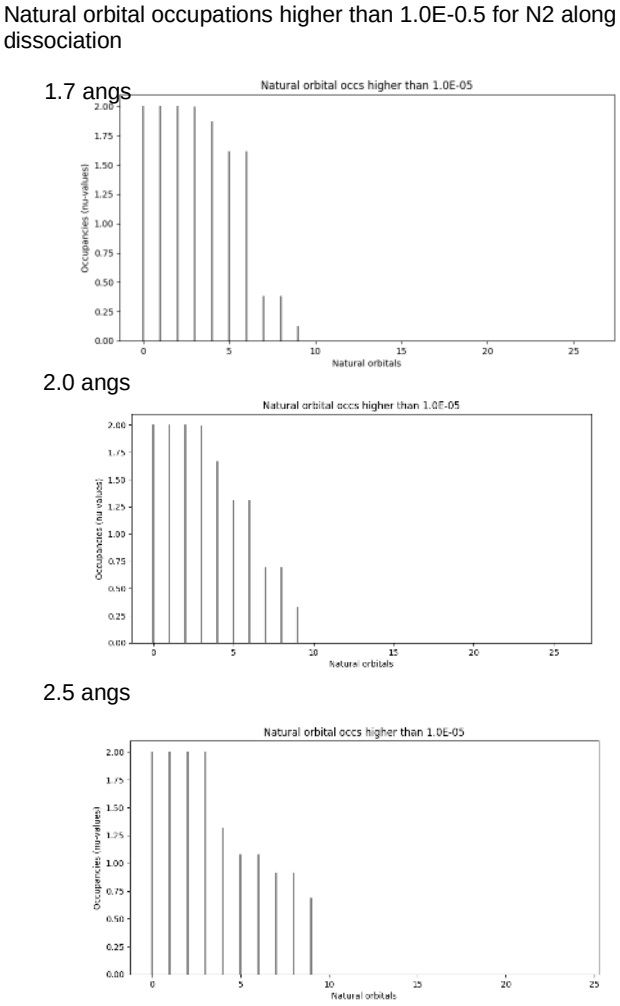


Figure 15: Natural orbital occupations (higher than a 1.0E-05) at different separations.

Potential energy surfaces for N_2

In Fig. 17, the potential energy surfaces computed within the 1S, single NOON scheme are depicted for different values used for the threshold parameter (ν), it is particularly interesting the effect of using a large threshold since results are largely affected giving wrong PES, while the proper NO space splitting seems to alleviate the problem, in bot cases the best choice resembles in the consideration and use of small NOONs. In Fig. 18, a bigger comparison is shown involving the rest of schemes.

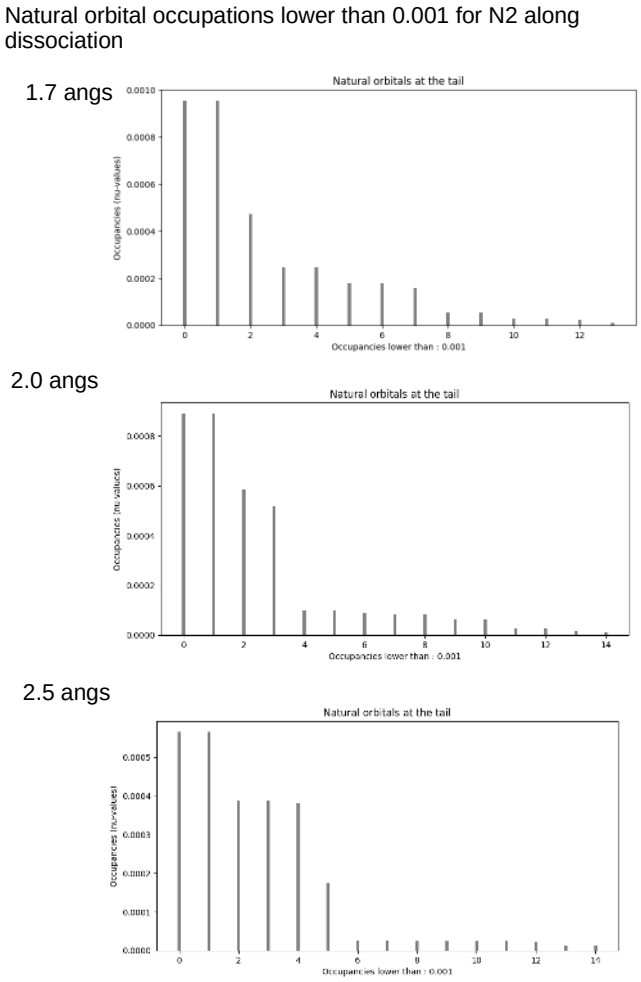


Figure 16: Natural orbital occupations (lower than a threshold of 0.001) at different separations.

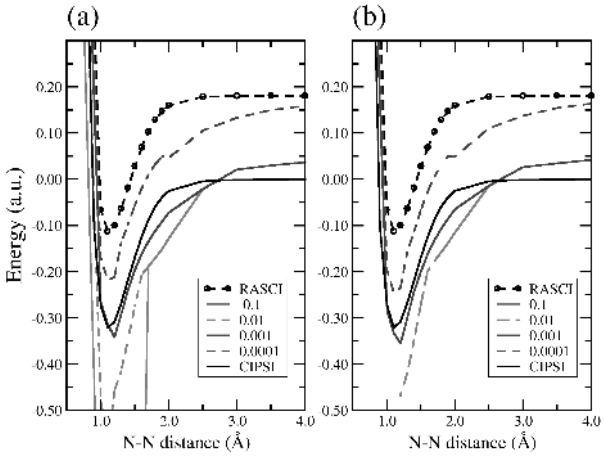


Figure 17: Absolute energy profiles along the N_2 dissociation computed at the RASCI (dashed black) and CIPSI (solid black) computed with the cc-pVDZ basis set and their comparison with (a) *on-top* DFT calculations and (b) splitting the natural orbital space. In both cases the 1S single NOON based threshold scheme was used (number in label resembles the occupation threshold used).

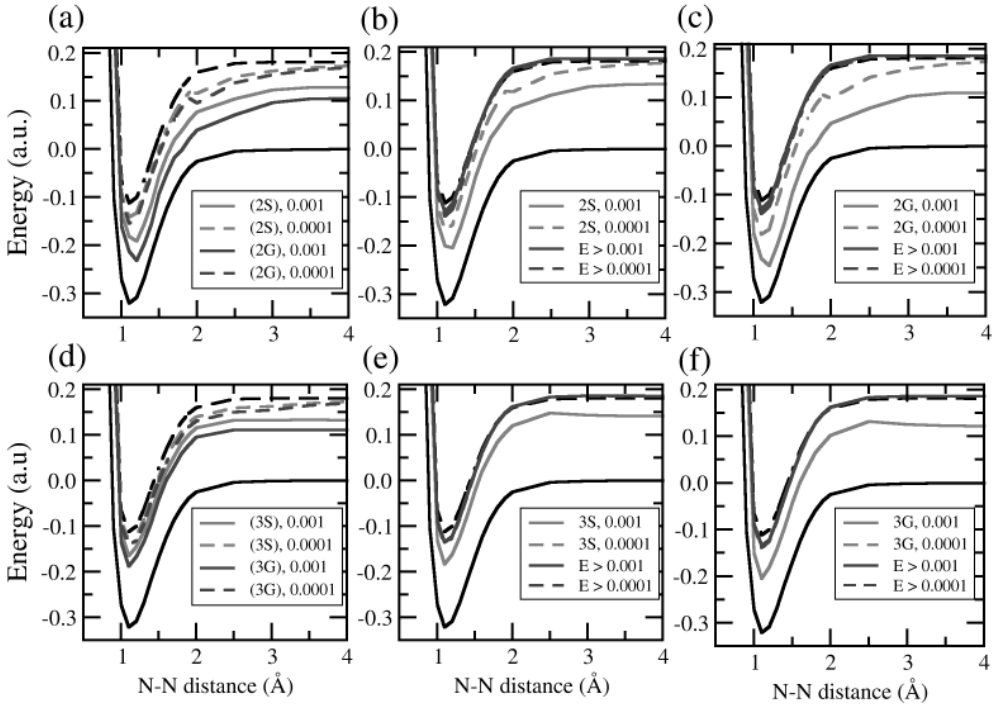


Figure 18: General comparison of N_2 PES computed with the different schemes presented in this work and standard RASCI and CIPSI calculations within the cc-pVDZ basis set. The $E_>$ label indicates such cases where no DFT correlation energy is incorpored into the final energy, i.e. these energies correspond to the WFT energy computed in the space of NOs with large occupations.

Table 15: Correlation energies (computed with *so*VWN(S) functional) and E_ν values computed within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold.

	Threshold values $E_{>\nu}$				Threshold values E_c-DFT			
d (Å)	0.1	0.01	0.001	0.0001	0.1	0.01	0.001	0.0001
0.5	0.0000	0.0000	-103.0837	-103.0876	-0.1493	-0.1022	-0.0425	0.0000
0.7	0.0000	-107.6549	-107.6852	-107.6890	-0.1936	-0.1193	-0.0447	0.0000
1.0	0.0000	-109.0378	-109.0486	-109.0521	-0.1321	-0.0915	-0.0431	0.0000
1.1	0.0000	-109.0849	-109.0950	-109.0997	-0.1446	-0.0913	-0.0484	0.0000
1.2	0.0000	-109.0702	-109.0825	-109.0852	-0.1120	-0.0913	-0.0400	0.0000
1.3	0.0000	-109.0282	-109.0406	-109.0431	-0.1119	-0.0923	-0.0431	0.0000
1.4	0.0000	-108.9776	-108.9901	-108.9924	-0.1109	-0.0951	-0.0462	0.0000
1.5	0.0000	-108.9275	-108.9415	-108.9425	-0.1111	-0.0994	-0.0420	0.0000
1.6	0.0000	-108.8827	-108.8973	-108.8981	-0.1126	-0.1041	-0.0417	0.0000
1.7	-108.8459	-108.8459	-108.8606	-108.8614	-0.1043	-0.1043	-0.0456	0.0000
1.8	-108.8183	-108.8183	-108.8323	-108.8329	-0.1033	-0.1033	-0.0436	0.0000
1.9	-108.7995	-108.7995	-108.8119	-108.8123	-0.1006	-0.1006	-0.0448	0.0000
2.0	-108.7878	-108.7878	-108.7981	-108.7984	-0.0961	-0.0961	-0.0416	0.0000
2.5	-108.7744	-108.7744	-108.7765	-108.7770	-0.0680	-0.0680	-0.0358	0.0000
3.0	-108.7740	-108.7740	-108.7742	-108.7747	-0.0529	-0.0529	-0.0422	0.0000
3.5	-108.7742	-108.7742	-108.7743	-108.7746	-0.0480	-0.0480	-0.0439	0.0000
4.0	-108.7744	-108.7744	-108.7744	-108.7747	-0.0481	-0.0481	-0.0441	0.0000

Table 16: RAS-soVWN(S) total energies and comparison with CASSCF and FCI calculations within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold. Dotted empty values cannot be computed since the threshold values pushes to cut the RAS2 space.

				Threshold values			
				Total RAs-soVWN(S) energy			
d (Å)	CIPSI	CASSCF(6,6)	RASCI	0.1	0.01	0.001	0.0001
0.5		0.0000	-103.0837	...	...	-103.1262	-103.0876
0.7	0.0000	-107.6549	-107.6852	...	-107.7741	-107.7300	-107.6890
1.0	0.0000	-109.0378	-109.0486	...	-109.1293	-109.0916	-109.0521
1.1	0.0000	-109.0849	-109.0950	...	-109.1762	-109.1434	-109.0997
1.2	0.0000	-109.0702	-109.0825	...	-109.1616	-109.1225	-109.0852
1.3	0.0000	-109.0282	-109.0406	...	-109.1205	-109.0838	-109.0431
1.4	0.0000	-108.9776	-108.9901	...	-109.0728	-109.0363	-108.9924
1.5	0.0000	-108.9275	-108.9415	...	-109.0269	-108.9834	-108.9425
1.6	0.0000	-108.8827	-108.8973	...	-108.9868	-108.9390	-108.8981
1.7	-108.8459	-108.8459	-108.8606	-108.9502	-108.9502	-108.9061	-108.8614
1.8	-108.8183	-108.8183	-108.8323	-108.9216	-108.9216	-108.8760	-108.8329
1.9	-108.7995	-108.7995	-108.8119	-108.9000	-108.9000	-108.8567	-108.8123
2.0	-108.7878	-108.7878	-108.7981	-108.8838	-108.8838	-108.8397	-108.7984
2.5	-108.7744	-108.7744	-108.7765	-108.8423	-108.8423	-108.8123	-108.7770
3.0	-108.7740	-108.7740	-108.7742	-108.8269	-108.8269	-108.8164	-108.7747
3.5	-108.7742	-108.7742	-108.7743	-108.8223	-108.8223	-108.8182	-108.7746
4.0	-108.7744	-108.7744	-108.7744	-108.8224	-108.8224	-108.8185	-108.7747

Table 17: Correlation energies (computed with *so*VWN(G) functional) and E_ν values computed within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold.

	Threshold values				Threshold values			
	$E_{>\nu}$				E_c-DFT			
d (Å)	0.1	0.01	0.001	0.0001	0.1	0.01	0.001	0.0001
0.5	0.0000	0.0000	-103.0837	-103.0876	-0.2147	-0.1470	-0.0611	0.0000
0.7	0.0000	-107.6549	-107.6852	-107.6890	-0.2792	-0.1720	-0.0645	0.0000
1.0	0.0000	-109.0378	-109.0486	-109.0521	-0.1910	-0.1323	-0.0623	0.0000
1.1	0.0000	-109.0849	-109.0950	-109.0997	-0.2093	-0.1321	-0.0700	0.0000
1.2	0.0000	-109.0702	-109.0825	-109.0852	-0.1621	-0.1322	-0.0579	0.0000
1.3	0.0000	-109.0282	-109.0406	-109.0431	-0.1621	-0.1337	-0.0625	0.0000
1.4	0.0000	-108.9776	-108.9901	-108.9924	-0.1607	-0.1378	-0.0669	0.0000
1.5	0.0000	-108.9275	-108.9415	-108.9425	-0.1610	-0.1440	-0.0608	0.0000
1.6	0.0000	-108.8827	-108.8973	-108.8981	-0.1632	-0.1508	-0.0604	0.0000
1.7	-108.8459	-108.8459	-108.8606	-108.8614	-0.1511	-0.1511	-0.0660	0.0000
1.8	-108.8183	-108.8183	-108.8323	-108.8329	-0.1497	-0.1497	-0.0632	0.0000
1.9	-108.7995	-108.7995	-108.8119	-108.8123	-0.1457	-0.1457	-0.0649	0.0000
2.0	-108.7878	-108.7878	-108.7981	-108.7984	-0.1392	-0.1392	-0.0603	0.0000
2.5	-108.7744	-108.7744	-108.7765	-108.7770	-0.0985	-0.0985	-0.0519	0.0000
3.0	-108.7740	-108.7740	-108.7742	-108.7747	-0.0767	-0.0767	-0.0611	0.0000
3.5	-108.7742	-108.7742	-108.7743	-108.7746	-0.0696	-0.0696	-0.0636	0.0000
4.0	-108.7744	-108.7744	-108.7744	-108.7747	-0.0696	-0.0696	-0.0639	0.0000

Table 18: RAS-*so*VWN(G) total energies and comparison with CASSCF and FCI calculations within the cc-pVDZ basis set, different values of the occupation threshold are indicated. All RAS3 data corresponds to the calculation made by cutting the NO space at the end of RAS2 and ussing all occupations in RAS3. All data is in atomic units, E_c-DFT is the correlation energy delivered by the density functional and $E_{>\nu}$ is the WFT energy computed with natural orbitals with occupation numbers larger than threshold. Dotted empty values cannot be computed since the threshold values pushes to cut the RAS2 space.

				Threshold values			
				Total RAs-*so*VWN(G) energy			
d (Å)	CIPSI	CASSCF(6,6)	RASCI	0.1	0.01	0.001	0.0001
0.5	-103.2506	-103.1412	-103.0698	...	....	-103.1448	-103.0876
0.7	-107.8521	-107.6829	-107.6681	...	-107.8572	-107.7498	-107.6890
1.0	-109.2324	-109.0449	-109.0257	...	-109.1809	-109.1109	-109.0521
1.1	-109.2811	-109.0908	-109.0722	...	-109.2271	-109.1650	-109.0997
1.2	-109.2686	-109.0757	-109.0601	...	-109.2148	-109.1405	-109.0852
1.3	-109.2282	-109.0330	-109.0234	...	-109.1743	-109.1031	-109.0431
1.4	-109.1782	-108.9810	-108.9781	...	-109.1279	-109.0570	-108.9924
1.5	-109.1286	-108.9296	-108.9324	...	-109.0854	-109.0023	-108.9425
1.6	-109.0842	-108.8841	-108.8912	...	-109.0481	-108.9577	-108.8981
1.7	-109.0476	-108.8472	-108.8571	-108.9971	-109.0117	-108.9266	-108.8614
1.8	-109.0192	-108.8197	-108.8311	-108.9680	-108.9820	-108.8955	-108.8329
1.9	-108.9988	-108.8011	-108.8125	-108.9452	-108.9576	-108.8768	-108.8123
2.0	-108.9851	-108.7899	-108.8002	-108.9270	-108.9372	-108.8583	-108.7984
2.5	-108.9640	-108.7777	-108.7813	-108.8729	-108.8750	-108.8284	-108.7770
3.0	-108.9614	-108.7754	-108.7792	-108.8507	-108.8509	-108.8353	-108.7747
3.5	-108.9605	-108.7769	-108.7791	-108.8439	-108.8439	-108.8379	-108.7746
4.0	-108.9603	-108.7770	-108.7793	-108.8440	-108.8441	-108.8383	-108.7747

Table 19: *On-Top* PES total energy results computed within the soVWN(S) and soVWN(G) functionals, the cc-pVDZ basis set and different threshold values as indicated.

	Threshold values RAS(soVWN(S))			
d (Å)	0.1	0.01	0.001	0.0001
0.5	-103.0698	-103.2489	-103.1844	-103.1203
0.7	-107.6681	-107.9017	-107.7960	-107.7192
1.0	-109.0257	-109.2721	-109.1343	-109.0733
1.1	-109.0722	-109.3469	-109.1824	-109.1249
1.2	-109.0601	-109.2243	-109.1721	-109.1060
1.3	-109.0234	-109.2145	-109.1353	-109.0724
1.4	-108.9781	-109.1995	-109.0891	-109.0300
1.5	-108.9324	-109.1851	-109.0435	-108.9806
1.6	-108.8912	-109.1774	-109.0038	-108.9383
1.7	-108.8571	-108.9713	-108.9713	-108.9080
1.8	-108.8311	-108.9452	-108.9452	-108.8789
1.9	-108.8125	-108.9242	-108.9242	-108.8604
2.0	-108.8002	-108.9069	-108.9069	-108.8444
2.5	-108.7813	-108.8556	-108.8556	-108.8277
3.0	-108.7792	-108.8370	-108.8370	-108.8272
3.5	-108.7791	-108.8320	-108.8320	-108.8271
4.0	-108.7793	-108.8315	-108.8315	-108.8273
	Threshold values RAS(soVWN(G))			
d (Å)	0.1	0.01	0.001	0.0001
0.5	-103.3274	-103.2347	-103.1425	-103.1025
0.7	-108.0049	-107.8526	-107.7419	-107.7015
1.0	-109.3820	-109.1828	-109.0946	-109.0567
1.1	-109.4697	-109.2317	-109.1485	-109.1015
1.2	-109.2978	-109.2222	-109.1265	-109.0920
1.3	-109.3002	-109.1855	-109.0944	-109.0542
1.4	-109.2987	-109.1388	-109.0533	-109.0059
1.5	-109.2985	-109.0934	-109.0022	-108.9612
1.6	-109.3060	-109.0544	-108.9595	-108.9209
1.7	-109.0225	-109.0225	-108.9308	-108.8876
1.8	-108.9965	-108.9965	-108.9005	-108.8618
1.9	-108.9744	-108.9744	-108.8819	-108.8429
2.0	-108.9549	-108.9549	-108.8642	-108.8295
2.5	-108.8890	-108.8890	-108.8486	-108.8102
3.0	-108.8630	-108.8630	-108.8488	-108.8052
3.5	-108.8557	-108.8557	-108.8487	-108.7956
4.0	-108.8550	-108.8550	-108.8489	-108.7904

Table 20: Selected RAS-*so*DFT results PES total energy results computed within the *so*VWN(S) and *so*VWN(G) functionals (the threshold value was set to 0.001 in both cases), the cc-pVTZ basis set and different threshold values as indicated.

d (Å)	CIPSI	RASCI	CASSCF(6,6)	RAS-*so*VWN(S)	RAS-*so*VWN(G)
0.5	-103.8083	-103.4813	-103.4725	-103.5428	-103.5697
0.7	-108.1109	-107.8040	-107.7960	-107.8623	-107.8881
0.9	-109.2309	-108.9113	-108.8905	-108.9692	-108.9949
1.0	-109.3714	-109.0588	-109.0830	-109.1163	-109.1420
1.1	-109.4004	-109.0964	-109.1194	-109.1552	-109.1816
1.2	-109.3762	-109.0784	-109.1001	-109.1360	-109.1618
1.3	-109.3303	-109.0374	-109.0557	-109.0936	-109.1188
1.4	-109.2790	-108.9922	-109.0030	-109.0459	-109.0700
1.5	-109.2302	-108.9495	-108.9513	-109.0047	-109.0295
1.6	-109.1871	-108.9105	-108.9054	-108.9687	-108.9949
1.7	-109.1508	-108.8772	-108.8679	-108.9355	-108.9617
1.8	-109.1210	-108.8509	-108.8397	-108.9082	-108.9339
1.9	-109.0970	-108.8317	-108.8204	-108.8870	-108.9119
2.0	-109.0779	-108.8185	-108.8084	-108.8741	-108.8990
2.5	-109.0292	-108.7976	-108.7949	-108.8528	-108.8776
3.0	-109.0157	-108.7956	-108.7938	-108.8509	-108.8758
3.5	-109.0121	-108.7956	-108.7946	-108.8332	-108.8501
4.0	-109.0108	-108.7960	-108.7947	-108.8210	-108.8322

Energy dependence on ν parameter

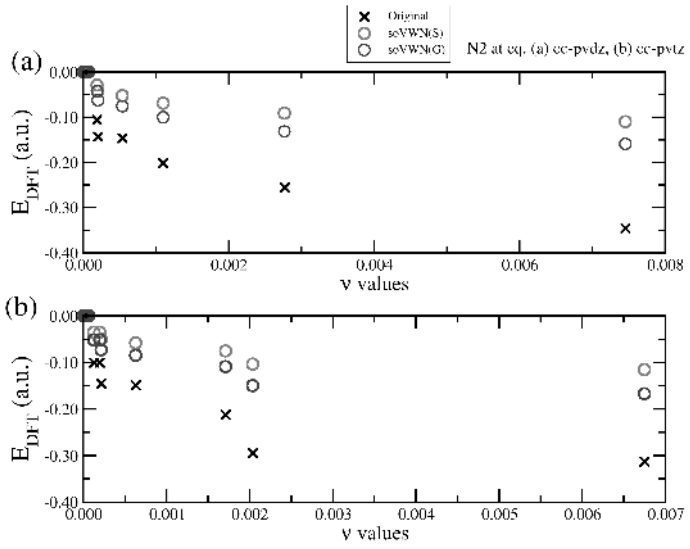


Figure 19: DFT energy dependence on the parameter nu for the N_2 ground state at $r_{NN} = 1.10$ Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets.

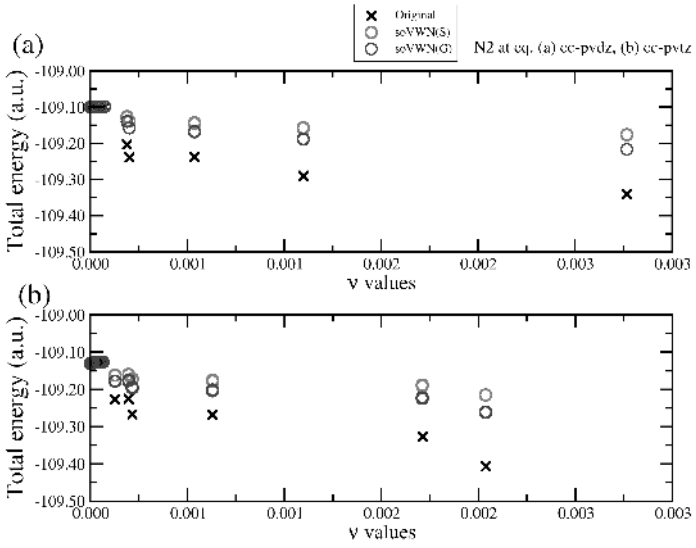


Figure 20: Total energy dependence on the parameter nu for the N_2 ground state at $r_{NN} = 1.10$ Å computed with the (a) cc-pVDZ and (b) cc-pVTZ basis sets.

Reducing discontinuities

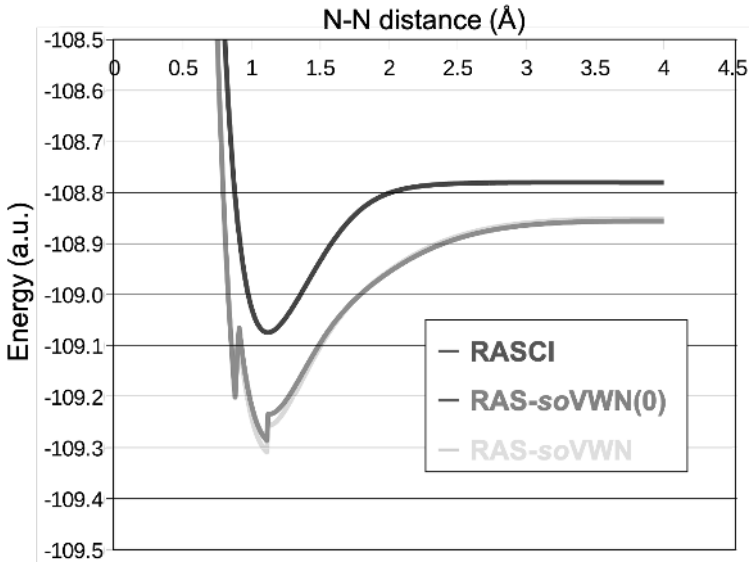


Figure 21: Absolute energy profiles along the N_2 dissociation computed at the RASCI (dark blue), RAS-soVWN (yellow), RAS-soVWN (orange) computed with the cc-pVDZ basis set. RAS-soVWN have been obtained by fixing the number of NOs entering the soVWN correlation to those corresponding whose sum of occupation numbers is lower than 0.001 at the equilibrium distance.

Discontinuities at larger basis sets

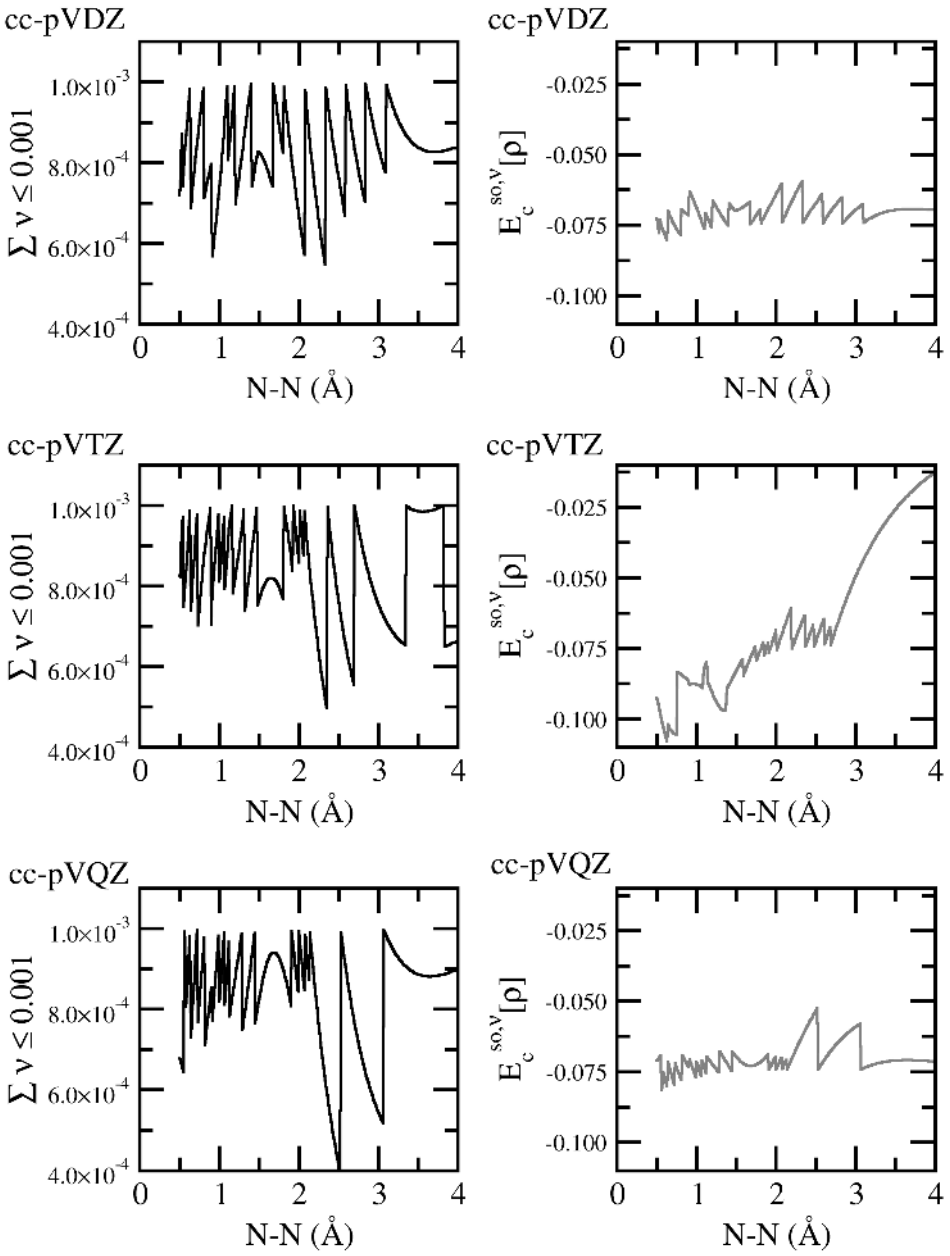


Figure 22: Sum of occupations lower than 0.001 and corresponding values of the DFT correlation functional *so*VWN(G). Results computed with different basis sets.

Size consistency analysis

Calculations for H_2 + nHe systems ($n = 1, 2, 3, 4$) with H $\cdots$ He distances equal to 100 Å. RAS-soVWN(G) calculations with RAS2 space with 2 electrons in 2 orbitals, and threshold = 0.001 for the soDFT correlation. Reference: ROHF triplet.

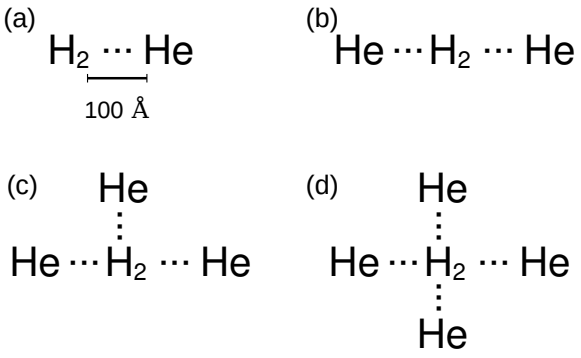


Figure 24: The different $H_2 \cdots n$He systems considered in this analysis. He atoms are placed orthogonally in the x- and y-axes, while H_2 bond lies on the z-axis.

Table 21: Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2 \cdots$ He system.

Basis set	ROHF	RAS-SF	soVWN(G)
Energy of total system (a.u.)			
cc-pVDZ	-3.6222	-4.0005	-4.0162
cc-pVTZ	-3.6355	-4.0111	-4.0278
cc-pVQZ	-3.6379	-4.0119	-4.0293
Sum of fragment energies (a.u.)			
cc-pVDZ	-3.6222	-4.0005	-4.0016
cc-pVTZ	-3.6355	-4.0111	-4.0120
cc-pVQZ	-3.6379	-4.0119	-4.0129
Absolute error (kcal/mol)			
cc-pVDZ	0.00	0.00	9.15
cc-pVTZ	0.00	0.00	9.91
cc-pVQZ	0.00	0.00	10.27

Table 22: Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2 \cdots 2\,He$ system.

Basis set	ROHF	RAS-SF	soVWN(G)
Energy of total system (a.u.)			
cc-pVDZ	-6.4774	-6.8557	-6.8763
cc-pVTZ	-6.4967	-6.8722	-6.8944
cc-pVQZ	-6.4994	-6.8734	-6.8964
Sum of fragment energies (a.u.)			
cc-pVDZ	-6.4774	-6.8556	-6.8568
cc-pVTZ	-6.4967	-6.8722	-6.8732
cc-pVQZ	-6.4994	-6.8734	-6.8744
Absolute error (kcal/mol)			
cc-pVDZ	0.00	0.04	12.25
cc-pVTZ	0.00	0.02	13.30
cc-pVQZ	0.01	0.03	13.77

Table 23: Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2 \cdots 3\,He$ system.

Basis set	ROHF	RAS-SF	soVWN(G)
Energy of total system (a.u.)			
cc-pVDZ	-9.3326	-9.7109	-9.7358
cc-pVTZ	-9.3578	-9.7334	-9.7602
cc-pVQZ	-9.3610	-9.7349	-9.7627
Sum of fragment energies (a.u.)			
cc-pVDZ	-9.3326	-9.7108	-9.7119
cc-pVTZ	-9.3578	-9.7334	-9.7343
cc-pVQZ	-9.3609	-9.7349	-9.7359
Absolute error (kcal/mol)			
cc-pVDZ	0.00	0.04	14.97
cc-pVTZ	0.01	0.02	16.24
cc-pVQZ	0.01	0.03	16.83

Table 24: Energies of the total system, sum of fragment energies and absolute error computed for the ROHF triplet references, RAS-SF and RAS-soVWN(G) singlet ground state of the $H_2 \cdots 4\,He$ system.

Basis set	ROHF	RAS-SF	soVWN(G)
Energy of total system (a.u.)			
cc-pVDZ	-12.1877	-12.5660	-12.5948
cc-pVTZ	-12.2190	-12.5945	-12.6255
cc-pVQZ	-12.2225	-12.5965	-12.6286
Sum of fragment energies (a.u.)			
cc-pVDZ	-12.1877	-12.5660	-12.5671
cc-pVTZ	-12.2190	-12.5945	-12.5955
cc-pVQZ	-12.2225	-12.5964	-12.5974
Absolute error (kcal/mol)			
cc-pVDZ	0.00	0.04	17.39
cc-pVTZ	0.01	0.02	18.87
cc-pVQZ	0.01	0.03	19.57

Appendix C

Solution for Attenuated Electron Repulsion Integrals Damped by a Gaussian Two-electron Operator

The derivation of the working equations for the two-electron repulsion integrals attenuated by a Gaussian operator (here referred as gauERIs) is provided. Obtained equations are suitable for a computational implementation of these integrals in the basis of Gaussian type orbitals. The following are some of the key points that drive the interest in this type of integrals during the development of the projects in this thesis.

- Most of the integral evaluation schemes are supported by recurrence relations, which starting from the fundamental electron repulsion integral (ERI, or a variation of it) can perform calculations for ERIs with higher angular momentum.[488]
- Among these schemes, the DRK method, based on the use of Gaussian quadrature,[418, 489] can not be used for the calculation of the attenuated Gaussian ERIs (gauERIs). This is due to the mathematical properties of the Rys polynomials and inconsistencies with the recurrence relations involved in the method. The Gaussian component introduces a shift in the convergence radius of the working kernel function, thus Rys polynomials cannot be used for the numerical computation of the gauERIs.
- Nevertheless, methods that use explicitly the complete form of the integrals, through the evaluation of Boys function are suitable for the calculation of these attenuated ERIs. From this observation, the derivation provided here conduces directly to the result shown by Boys.[102, 103] The following corresponds to the generalization of the result provided by Boys for arbitrary angular momentum basis sets.
- Current implementation of a code able for the calculation of these ERIs is based on the direct computation of the integrals as presented in this appendix.

Solution for the Gaussian two-electron repulsion integrals in the Boys picture.

The integrals in which we are interested are of the form:

$$\left(AB\middle|\frac{\exp\left(-\nu r_{ij}^2\right)}{r_{if}}\middle|CD\right) \tag{1}$$

in which ν is an arbitrary parameter ($\nu \in \mathbb{R}$), while A, B, C and D represent arbitrary contracted Gaussian type orbitals centered at nucleus **A**, **B**, **C** and **D**, respectively. The

operator appearing in equation 1 can be considered as the Coulomb electrostatic potential, attenuated by a Gaussian factor that acts over electrons i and j. In order to compute this, we wish to form a class of contracted integrals, first, we note that each basis function is k-fold contracted, thus, each contracted integral can be expressed as a sum of its component primitive integrals, which in turn are computed individually, i.e.

$$\Big(AB\Big|\frac{\exp\big(-\nu r_{12}^2\big)}{r_{12}}\Big|CD\Big)= \sum_{a=1}^{k_a}\sum_{b=1}^{k_b}\sum_{c=1}^{k_c}\sum_{d=1}^{k_d} d_a d_b d_c d_d N_a N_b N_c N_d \Big[AB\Big|\frac{\exp\Big(-\nu r_{ij}^2\Big)}{r_{ij}}\Big|CD\Big] \tag{2}$$

where d_a, d_b, ..., are the contraction coefficients and N_a, N_b, ... are the normalization coefficients for the contracted Gaussian functions. The problem arises for the evaluation of the contracted integral

$$G \equiv \Big[AB\Big|\frac{\exp\Big(-\nu r_{ij}^2\Big)}{r_{ij}}\Big|CD\Big]= \iint \chi_a\chi_b \frac{\exp\Big(-\nu r_{ij}^2\Big)}{r_{ij}} \chi_c\chi_d \, dV_i \, dV_j \tag{3}$$

where $dV_i = dx_i \, dy_i \, dz_i$ if the differential volume for the i-electron coordinates, and each χ_a, ..., are primitive contracted functions of the form:

$$\chi_a = x_{Ai}^{l_A} \, y_{Ai}^{m_A} \, z_{Ai}^{n_A} \, \exp\big(-\alpha_a r_{Ai}^2\big) \tag{4}$$

Being angular momentum quantum numbers l, m, n in the x, y ,z directions respectively and primitive exponent α_a. Here I will provide a way to solve the integral 3, due to the coupled nature of the Gaussian operator, the integration involve the use of more expansions than those appearing for the standard ERIs ($\big[AB\big|CD\big]$). The key step ingredient for the integration is the treatment of electron-i components as a static variable while electron-j components are integrated, this is mandatory due to the coupling given by the exponential factor ($-\nu r_{ij}^2$). Essentially we can perform the integration by separating the Cartesian components of G leading to similar equations for each component.

We can write G explicitly as

$$\begin{aligned} G = \iint & x_{Ai}^{l_A} \, y_{Ai}^{m_A} \, z_{Ai}^{n_A} \, \exp\big(-\alpha_a r_{Ai}^2\big) x_{Bi}^{l_B} \, y_{Bi}^{m_B} \, z_{Bi}^{n_B} \, \exp\big(-\alpha_b r_{Bi}^2\big) \\ & \frac{\exp\Big(-\nu r_{ij}^2\Big)}{r_{ij}} x_{Cj}^{l_C} \, y_{Cj}^{m_C} \, z_{Cj}^{n_C} \, \exp\big(-\alpha_c r_{Cj}^2\big) x_{Dj}^{l_D} \, y_{Dj}^{m_D} \, z_{Dj}^{n_D} \, \exp\big(-\alpha_d r_{Dj}^2\big) \, dV_i \, dV_j \end{aligned} \tag{5}$$

where $\vec{\boldsymbol{r_{Ai}}} = \vec{\boldsymbol{r_i}} - \vec{\boldsymbol{A}}$ the distance between the ith-electron and the fixed center (nucleus A), and $\vec{\boldsymbol{x_{Ai}}} = \vec{\boldsymbol{x_i}} - \vec{\boldsymbol{A}}_{\boldsymbol{x}}$ are the x-components of $\vec{\boldsymbol{r_{Ai}}}$, and so on. The expression 5 is noting but the Cartesian separation of the angular components of G. First we will treat each of these terms in a separate way.

Treatment of the Cartesian angular momentum components

The treatment of the angular components can be done in terms of the reduced centers defined as

$$\begin{aligned}\vec{\boldsymbol{P}} &= \frac{\alpha_a \vec{\boldsymbol{A}} + \alpha_b \vec{\boldsymbol{B}}}{\alpha_a + \alpha_b} \\ \vec{\boldsymbol{Q}} &= \frac{\alpha_c \vec{\boldsymbol{C}} + \alpha_d \vec{\boldsymbol{D}}}{\alpha_c + \alpha_d}\end{aligned} \tag{6}$$

From here on, we will focus, most of the time on one of the Cartesian components involved in the expression for G (equation 5), since the rest of the components can be treated in a similar way. Particularly, the product of two given angular components of the form $x^{l_A} x^{l_B}$ can be written as:

$$x^{l_A} x^{l_B} = (x - A_x)^{l_A} (x - B_x)^{l_B} \tag{7}$$

where A_x and B_x are the x-cartesian components of the nuclear coordinates A and B. In terms of the x-components of $\vec{\boldsymbol{P}}$, we have:

$$\begin{aligned}x^{l_A} x^{l_B} &= \big\{(x - P_x) + (P_x - A_x)\big\}^{l_A} \big\{(x - P_x) + (P_x - B_x)\big\}^{l_B} \\ &= (x_P + PA_x)^{l_A} (x_P + PB_x)^{l_B}\end{aligned} \tag{8}$$

where PA_x and PB_x are the x-components of the vectors $\vec{\boldsymbol{PA}}$ and $\vec{\boldsymbol{PB}}$, and x_P is the x-component of the $\vec{\boldsymbol{xP}}$ vector. We can expand the product by a binomial expansion and get

$$x^{l_A} x^{l_B} = \sum_{i=0}^{l_A} \sum_{j=0}^{l_B} \binom{l_A}{i} \binom{l_B}{j} x_P^i x_P^j PA_x^{l_A - i} PB_x^{l_B - j} \tag{9}$$

From this expression, and noting that for each element k, such that $i + j = k$, we can form products of the type

$$\begin{aligned}k = i + j = 0; &\binom{l_A}{0} \binom{l_B}{0} PA_x^{l_A} PB_x^{l_B} \\ k = i + j = 1; &\binom{l_A}{0} \binom{l_B}{1} PA_x^{l_A} PB_x^{l_B - 1} \\ &\binom{l_A}{1} \binom{l_B}{0} PA_x^{l_A - 1} PB_x^{l_B} \\ &\vdots\end{aligned} \tag{10}$$

we can define the function

$$C_k(l, m, a, b) = \sum_{i=0}^{l} \sum_{j=0}^{m} \binom{l}{i} \binom{m}{j} a^{l-i} b^{m-j} \tag{11}$$

which, under the condition $i + j = k$, the following expression is true

$$\sum_{k=0}^{l_A+l_B} C_k(l_A, l_B, a, b) = C_0 + C_1 + C_2 + \ldots + C_{l_A} + C_{l_B} \tag{12}$$

This is a way to organize in a condensed function C_k all the products in equation 9. Thus if and only if the condition $i + j = k$ is satisfied, the product of angular components can be written as

$$x^{l_A} x^{l_B} = \sum_{l_P=0}^{l_A+l_B} C_{l_P}(l_A, l_B, PA_x, PB_x) x_P^{l_P} \tag{13}$$

And analogously for the y- and z-components

$$y^{m_A} y^{m_B} = \sum_{m_P=0}^{m_A+m_B} C_{m_P}(m_A, m_B, PA_y, PB_y) x_P^{l_P} \tag{14}$$

$$z^{n_A} z^{n_B} = \sum_{n_P=0}^{n_A+n_B} C_{n_P}(n_A, n_B, PA_z, PB_z) x_P^{n_P} \tag{15}$$

These expressions hold for the i-th components in G, for the j-th components (associated to $\vec{\boldsymbol{Q}}$) the treatment is analogous. We can rewrite 5 as

$$\begin{aligned} G = & \sum_{l_P=0}^{l_A+l_B} \sum_{m_P=0}^{m_A+m_B} \sum_{n_P=0}^{n_A+n_B} C_{l_P} C_{m_P} C_{n_P} \sum_{l_Q=0}^{l_C+l_D} \sum_{m_Q=0}^{m_C+m_D} \sum_{n_Q=0}^{n_C+n_D} C_{l_Q} C_{m_Q} C_{n_Q} \\ & \iint dV_i \, dV_j x_P^{l_P} y_P^{m_P} z_P^{n_P} \exp\left(-\alpha_a r_{Ai}^2\right) \exp\left(-\alpha_b r_{Bi}^2\right) \left(\frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}} \right) \\ & x_Q^{l_Q} y_Q^{m_Q} z_Q^{n_Q} \exp\left(-\alpha_c r_{Cj}^2\right) \exp\left(-\alpha_d r_{Dj}^2\right) \end{aligned} \tag{16}$$

With this treatment we have changed the information from labels i, j to labels P and Q, at least for the angular part an integration over i-electron coordinates involves an integration over P-coordinates and an integration over j-electron coordinates involves an integration over Q-coordinates.

Use of the Gauss product theorem

The inner integrals that appear in equation 16 involve the product of five Gaussian functions. We can apply the Gauss product theorem in order to get new functions that depend on the centers $\vec{\boldsymbol{P}}$ and $\vec{\boldsymbol{Q}}$ by considering this, the new expression for G is:

$$G = K_{AB}K_{CD}\sum_{l_P=0}^{l_A+l_B}\sum_{m_P=0}^{m_A+m_B}\sum_{n_P=0}^{n_A+n_B} C_{l_P}C_{m_P}C_{n_P}\sum_{l_Q=0}^{l_C+l_D}\sum_{m_Q=0}^{m_C+m_D}\sum_{n_Q=0}^{n_C+n_D} C_{l_Q}C_{m_Q}C_{n_Q}$$
$$\iint dV_i\, dV_j x_P^{l_P} y_P^{m_P} z_P^{n_P} \exp\left(-\gamma_P r_{Pi}^2\right)\left(\frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}}\right) x_Q^{l_Q} y_Q^{m_Q} z_Q^{n_Q} \exp\left(-\gamma_Q r_{Qj}^2\right) \tag{17}$$

where $dV_i = dx_P dy_P dz_P \;\; dV_j = dx_Q dy_Q dz_Q$; $\gamma_P = \alpha_a + \alpha_b$, $\gamma_Q = \alpha_c + \alpha_d$. And the two additional new factors K_{AB}, K_{CD} defined as:

$$\begin{aligned} K_{AB} &\equiv \exp\left(\frac{-\alpha_a\alpha_b}{\gamma_P}\|AB\|^2\right) \\ K_{CD} &\equiv \exp\left(\frac{-\alpha_c\alpha_d}{\gamma_Q}\|CD\|^2\right) \end{aligned} \tag{18}$$

Solution of the inner integral (G_{RI}) I: Treatment of the attenuated Coulomb operator

In the following we will focus on the treatment of the inner integral from equation 17, which will be referred to as G_{RI}, again it is convenient to expand the r_{ij} operator in terms of $\vec{\boldsymbol{P}}$ and $\vec{\boldsymbol{Q}}$ as:

$$\vec{\boldsymbol{r}}_{ij} = \vec{\boldsymbol{r}}_{Pi} - \vec{\boldsymbol{r}}_{Qj} + \vec{\boldsymbol{PQ}} \tag{19}$$

We need to face with the more challenging r_{ij}^2 operator, which can be expanded as:

$$\vec{\boldsymbol{r}}_{ij}^2 = \vec{\boldsymbol{r}}_{Pi}^2 + \vec{\boldsymbol{r}}_{Qj}^2 + \vec{\boldsymbol{PQ}}^2 + 2\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{PQ}} - 2\vec{\boldsymbol{r}}_{Qj}\cdot\vec{\boldsymbol{PQ}} - 2\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{r}}_{Qj} \tag{20}$$

By applying the Fourier transform to the Coulomb operator, we can write $1/r_{ij}$ as

$$\frac{1}{r_{ij}} = \frac{1}{2\pi^2}\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^{\boldsymbol{2}}}\exp\left(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r_{ij}}}\right) \tag{21}$$

Which in terms of 19, the Coulomb operator can be expressed as:

$$\frac{1}{r_{ij}} = \frac{1}{2\pi^2}\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^{\boldsymbol{2}}}\exp\left(i\vec{\boldsymbol{k}}\cdot(\vec{\boldsymbol{r}}_{Pi} - \vec{\boldsymbol{r}}_{Qj} + \vec{\boldsymbol{PQ}})\right) \tag{22}$$

By expressing the Coulomb operator in this way and multiplying by $\exp\left(-\nu r_{ij}^2\right)$, we can express the attenuated operator in the two-electron integrals as

$$\frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}} = \frac{1}{2\pi^2}\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^{\boldsymbol{2}}}\exp\left(-\nu r_{ij}^2\right)\exp\left(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Pi}\right)\exp\left(-i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Qj}\right)\exp\left(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{PQ}}\right) \tag{23}$$

This expression will be useful for further treatment of the integrals in G_{RI} (equation17), the Gaussian-attenuated Coulomb operator is then expressed as an integral in K-space that involves the product of four exponential functions one of them depends exclusively on the $\vec{\boldsymbol{P}}$ and $\vec{\boldsymbol{Q}}$ coordinates. The rest of Gaussian functions depend on the electron coordinates i, j. The kernel of the integral that appears in G_{RI}, involves the product of three exponential functions, that, using the Fourier transform of the Coulomb operator (equation 23), can be expanded as:

$$\begin{aligned}\exp\left(-\gamma_P r_{Pi}^2\right)\left(\frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}}\right)\exp\left(-\gamma_Q r_{Qj}^2\right) = \frac{1}{2\pi^2}\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^{\boldsymbol{2}}}\exp\left(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{PQ}}\right)\\ \exp\left(-\gamma_P\vec{\boldsymbol{r}}_{Pi}^2 + i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Pi} - \gamma_Q\vec{\boldsymbol{r}}_{Qj}^2 - i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Qj}\right)\exp\left(-\nu r_{ij}^2\right).\end{aligned} \tag{24}$$

By further substitution of the equation 20, for r_{ij} in terms of the $\vec{\boldsymbol{P}}$ and $\vec{\boldsymbol{Q}}$ centers, we can re-write the equation 24 as

$$\begin{aligned}\exp\left(-\gamma_P r_{Pi}^2\right)\left(\frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}}\right)\exp\left(-\gamma_Q r_{Qj}^2\right) = \frac{1}{2\pi^2}\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^{\boldsymbol{2}}}\exp\left(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{PQ}}\right)\\ \exp\left(-\gamma_P\vec{\boldsymbol{r}}_{Pi}^2 + i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Pi} - \gamma_Q\vec{\boldsymbol{r}}_{Qj}^2 - i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Qj}\right)\\ \exp\left(-\nu\vec{\boldsymbol{r}}_{Pi}^2 - \nu\vec{\boldsymbol{r}}_{Qj}^2 - \nu\vec{\boldsymbol{PQ}}^2 - 2\nu\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{PQ}} + 2\nu\vec{\boldsymbol{r}}_{Qj}\cdot\vec{\boldsymbol{PQ}} + 2\nu\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{r}}_{Qj}\right).\end{aligned} \tag{25}$$

By multiplying the exponential functions and grouping similar terms we can write:

$$\begin{aligned}\exp\left(-\gamma_P r_{Pi}^2\right)\left(\frac{\exp\left(-\nu r_{ij}^2\right)}{r_{ij}}\right)\exp\left(-\gamma_Q r_{Qj}^2\right) = \frac{1}{2\pi^2}\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^{\boldsymbol{2}}}\exp\left(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{PQ}}\right)\\ \exp(2\nu\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{r}}_{QJ})\exp\left(-\nu\vec{\boldsymbol{PQ}}^2\right)\exp\left(-(\gamma_P+\nu)\vec{\boldsymbol{r}}_{Pi}^2 + i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Pi} - 2\nu\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{PQ}}\right)\\ \exp\left(-(\gamma_Q+\nu)\vec{\boldsymbol{r}}_{Qj}^2 + i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Qj} - 2\nu\vec{\boldsymbol{r}}_{Qj}\cdot\vec{\boldsymbol{PQ}}\right)\end{aligned} \tag{26}$$

Solution of the inner integral (G_{RI}) II: Solutions for the Cartesian components.

Within the previous considerations, G_{RI} becomes

$$\begin{aligned}G_{RI} = \frac{1}{2\pi^2}\exp\left(-\nu\vec{\boldsymbol{PQ}}^2\right)\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^{\boldsymbol{2}}}\iint dV_i\, dV_j\\ x^{l_P}y^{m_P}z^{n_P}x^{l_Q}y^{m_Q}z^{n_Q}\exp\left(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{PQ}}\right)\exp(2\nu\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{r}}_{Qj})\\ \exp\left(-(\gamma_P+\nu)\vec{\boldsymbol{r}}_{Pi}^2 + i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Pi} - 2\nu\vec{\boldsymbol{r}}_{Pi}\cdot\vec{\boldsymbol{PQ}}\right)\\ \exp\left(-(\gamma_Q+\nu)\vec{\boldsymbol{r}}_{Qj}^2 + i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{r}}_{Qj} - 2\nu\vec{\boldsymbol{r}}_{Qj}\cdot\vec{\boldsymbol{PQ}}\right)\end{aligned} \tag{27}$$

In fact, the introduction of $\vec{\boldsymbol{P}}$ and $\vec{\boldsymbol{Q}}$ vectors and further expression of r_{ij} in terms of them, allow us to separate this expression into three Cartesian components ($G_{RI}^{(x)}$,$G_{RI}^{(y)}$,$G_{RI}^{(z)}$).

$$G_{RI} = \frac{1}{2\pi^2}\exp\Big(-\nu\vec{\boldsymbol{PQ}}^2\Big)\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^2}G_{RI}^{(x)}G_{RI}^{(y)}G_{RI}^{(z)} \tag{28}$$

Where each component G_{RI}^{η} ($\eta = x, y, z$) have the same form given by (here the x-component is shown, the rest are analogous)

$$\begin{aligned} G_{RI}^{(x)} = \iint dx_P\, dx_Q\, x_P^{l_P}\, x_Q^{l_Q}\, \exp(2\nu x_P x_Q) \\ \exp\big((-\gamma_P+\nu)x_P^2\big) + (ik_x - 2\nu\overline{PQ}_x)x_P) \\ \exp\big((-\gamma_Q+\nu)x_Q^2\big) + (-ik_x + 2\nu\overline{PQ}_x)x_Q) \end{aligned} \tag{29}$$

These $G_{RI}^{(x,y,z)}$ integrals can be solved by further separation of the kernels into a x_P-dependent part and another that does not depend on this variable. To this end, first we note that integral 29 can be written as

$$\begin{aligned} G_{RI}^{(x)} = \int dx_P\, x_P^{l_P}\exp\big((-\gamma_P+\nu)x_P^2 + ik_x + 2\nu\overline{PQ}_x\big)x_P) \\ \int dx_Q\, x_Q^{l_Q}\,\exp\big((-\gamma_Q+\nu)x_Q^2 + (-ik_x + 2\nu(\overline{PQ}_x + x_P))x_Q\big) \\ = I_P x_P I_Q(x_Q; x_P) \end{aligned} \tag{30}$$

where the integration is expressed as the multiplication of two integrals dependent on the two x_P and x_Q variables. First Integration over x_Q (the x_P runs out of the integral) involves the solution of an integral of the form:

$$\begin{aligned} I_Q = \int dx_Q x_Q^{l_Q}\exp\big(-ax_Q^2 + (b-ic)x_Q\big) \\ = \exp\big((b-ic)^2/4a\big)\int dx_Q x_Q^{l_Q}\exp\Big(-a(x_Q - \frac{b-ic}{2a})^2\Big) \end{aligned} \tag{31}$$

with $a \equiv (\nu+\gamma_Q)$; $b \equiv 2\nu(x_P + \overline{PQ}_x))$; $c \equiv k_x$. By the change of variable $u \equiv x_Q - (b-ic)/2a$, we have te simpler expression

$$I_Q = \exp\big((b-ic)^2/4a\big)\int du\Big[u + \frac{b-ic}{2a}\Big]^{l_Q}\exp\big(-au^2\big) \tag{32}$$

for which, by expanding the binomial faction (always letting the integration variable with the simplest expansion exponent) we can write I_Q as:

$$I_Q = \exp\big((b-ic)^2/4a\big)\sum_{r=0}^{[l_Q]}\frac{l_Q!}{r!(l_Q-r)!}\left[\frac{(b-ic)}{(2a)}\right]^{l_Q-r}\int du\, u^r \exp\big(-au^2\big) \tag{33}$$

where r is just the index of the summation. The final Gaussian integral in 33 is only non-zero for even-r values, this condition affects the summation index and the whole integral reduces to

$$\begin{aligned} I_Q &= \sqrt{\frac{\pi}{a}}\exp\big((b-ic)^2/4a\big)\sum_{r=0}^{[l_Q/2]}\frac{l_Q!}{(2r)!(l_Q-2r)!}\left[\frac{(b-ic)}{(2a)}\right]^{l_Q-2r}\frac{(2r)!}{2^{2r}r!a^r} \\ I_Q &= \sqrt{\frac{\pi}{a}}\exp\left(\frac{(b-ic)^2}{4a}\right)\sum_{r=0}^{[l_Q/2]}\frac{l_Q!}{r!(l_Q-2r)!}\left[\frac{(b-ic)^{l_Q-2r}}{(2^{l_Q}\ a^{l_Q-r})}\right] \end{aligned} \tag{34}$$

where one can trivially check that the following relations hold

$$(b-ic)^2 = \left[2\nu(x_P+\overline{PQ}_x)-ik_x\right]^2$$

$$\frac{(b-ic)^2}{4a} = \frac{4\nu^2(x_P+\overline{PQ}_x)^2-4\nu ik_x(x_P+\overline{PQ}_x)-k_x^2}{4(\nu+\gamma_Q)}$$

with which 34 becomes

$$I_Q = \sqrt{\frac{\pi}{a}}\exp\left(\frac{(b-ic)^2}{4a}\right)\sum_{r=0}^{[l_Q/2]}\frac{l_Q!}{r!(l_Q-2r)!}\left[\frac{(2\nu(x_P+\overline{PQ}_x)-ik_x)^{l_Q-2r}}{(2^{l_Q}\ a^{l_Q-r})}\right] \tag{35}$$

The multiplicative exponential term can be separated into the x_P-dependent term and an x_P-dependent term. We want to make such separation since the next step will consist on the evaluation of the integral over x_P space. By expanding the argument in the exponential of expression 34:

$$\exp\left(\frac{(b-ic)^2}{4a}\right) = \exp\left(\frac{4\nu^2x_P^2+8\nu^2\overline{PQ}_x x_P-4\nu ik_x x_P}{4(\nu+\gamma_Q)}\right)\exp\left(\frac{4\nu^2\overline{PQ}_x^2-4\nu ik_x\overline{PQ}_x-k_x^2}{4(\nu+\gamma_Q)}\right) \tag{36}$$

The first exponential function constitutes the x_P-dependent term, while the second is independent on this variable, and consequently can be taken outside the integral as a multiplicative factor. Furthermore, integral I contains a binomial term that is convenient to expand in order to get a separation of x_P from $\overline{PQ}_x$. Alternatively, one can avoid such expansion and continue with the next integration by choosing a suitable variable. However, this is not a good idea since the final expression become too complicated, and not suitable for computational implementation. The binomial factor in 35

$$(2\nu(x_P+\overline{PQ}_x)-ik_x)^{l_Q-2r}=\sum_{s=0}^{l_Q-2r}\frac{(l_Q-2r)!}{s!(l_Q-2r-s)!}(-i)k_x^{l_Q-2r-s}(2\nu)^s(x_P+\overline{PQ}_x)^s$$

and by expanding the secondary binomial $(x_P+\overline{PQ}_x)^s$ one gets

$$\begin{aligned}(2\nu(x_P+\overline{PQ}_x)-ik_x)^{l_Q-2r}=\sum_{s=0}^{l_Q-2r}\sum_{t=0}^{s}\frac{(l_Q-2r)!(-1)^{l_Q-s}}{t!(s-t)!(l_Q-2r-s)!}\\(ik_x)^{l_Q-2r-s}k_x^{l_Q-2r-s}(2\nu)^s(\overline{PQ}_x)^{s-t}x_P^t\end{aligned}\tag{37}$$

The substitution of expression 37 and 36 into 35 constitutes the integration of 30 over x_Q variable. As a result, $G_{RI}^{(x)}$ is reduced to an integral over x_P multiplied by the corresponding previous expansions, this integral is of the form

$$I_P=\int dx_P\,(x_P)^{l_P+t}\exp\left(\frac{-\delta}{\eta}x_P^2+\frac{\gamma_Q}{\eta}(ik_x-2\nu\overline{PQ}_x)x_P\right)\tag{38}$$

where $\delta\equiv\nu\gamma_P+\nu\gamma_Q+\gamma_P\gamma_Q$. In order to express this integral in sums as before, we can progress following the same strategy as before, by reducing the exponential term to a similar form as integral 33, the change of variable $u=x_P-(ib-c)/2a$ with $a\equiv\frac{\delta}{\eta}$, $b\equiv\gamma_Q k_x/\eta$ and $c\equiv(2\nu\gamma_Q\overline{PQ}_x)/\eta$ leads to

$$\begin{aligned}I_P=\int dx_P\,(x_P)^{l_P+t}\exp\left(-ax_P^2+(ib-c)x_P\right)\\=\exp\left(\frac{(ib-c)^2}{4a}\right)\int du\left[u+\frac{(ib-c)}{2a}\right]^{l_P+t}\exp\left(-au^2\right)\end{aligned}\tag{39}$$

By expanding the binomial (using the α index in this case)

$$\left[u+\frac{(ib-c)}{2a}\right]^{l_P+t}=\sum_{\alpha=0}^{l_P+t}\frac{(l_P+t)!}{\alpha!(l_P+t-\alpha)!}\left[\frac{(ib-c)}{2a}\right]^{l_P+t-\alpha}u^\alpha$$

thus

$$I_P=\exp\left(\frac{(ib-c)^2}{4a}\right)\sum_{\alpha=0}^{l_P+t}\frac{(l_P+t)!}{\alpha!(l_P+t-\alpha)!}\left[\frac{(ib-c)}{2a}\right]^{l_P+t-\alpha}\int du\,u^\alpha\exp\left(-au^2\right)$$

for which the Gaussian integral is non-zero only for even members ($\alpha=2\alpha$), integration yields

$$I_P=\exp\left(\frac{(ib-c)^2}{4a}\right)\sqrt{\frac{\pi}{a}}\sum_{\alpha=0}^{\left[\frac{l_P+t}{2}\right]}\frac{(l_P+t)!}{\alpha!(l_P+t-2\alpha)!}\left[\frac{(ib-c)^{l_P+t-2\alpha}}{(2a)^{l_P+t}}\right]a^\alpha$$

by switching $a\,b\,c$ to the original variables $\eta\,\delta\,\gamma_Q$, etc. We obtain the final expression

$$I_P = \sum_{\alpha=0}^{\left[\frac{l_P+t}{2}\right]} \frac{(l_P+t)!}{\alpha!(l_P+t-2\alpha)!}\sqrt{\frac{\pi\eta}{\delta}}\exp\left(\frac{\gamma_Q^2(ik_x-2\nu\overline{PQ}_x)^2}{4\eta\delta}\right)\left[\frac{\eta^\alpha(\gamma_Q(ik_x-2\nu\overline{PQ}_x))^{l_P+t-2\alpha}}{2^{l_P+t}\delta^{l_P+t-\alpha}}\right] \tag{40}$$

Substitution of integrals I_P and I_Q into 30, lead to the complete form of $G_{RI}^{(x)}$, expressed as an expansion over four indexes

$$\begin{aligned} G_{RI}^{(x)} = &\sum_{r=0}^{\left[\frac{l_Q}{2}\right]}\sum_{s=0}^{l_Q-2r}\sum_{t=0}^{s}\sum_{\alpha=0}^{\left[\frac{l_P+t}{2}\right]}\frac{l_Q!\,(l_P+t)!}{r!(s-t)!\,t!\,(l_P+t-2\alpha)!\,\alpha!} \\ &(-1)^{l_Q-s}\,(i)^{l_Q-2r-s}\left(\frac{1}{2^{l_Q}\eta^{l_Q-r}}\right)(2\nu)^s(\overline{PQ}_x)^{s-t}k_x^{l_Q-2r-s}\,\frac{\pi}{\sqrt{\delta}} \\ &\exp\left(\frac{\gamma_P+\gamma_Q}{4\delta}(ik_x-2\nu\overline{PQ}_x)^2\right)\frac{\eta^\alpha\left[\gamma_Q(ik_x-2\nu\overline{PQ}_x)\right]^{l_P+t-2\alpha}}{2^{l_P+t}\delta^{l_P+t-2\alpha}} \end{aligned} \tag{41}$$

Further expansions in G_{RI} and k-space integration

Expression 41 corresponds to the x-component of the overall G_{RI} integral (28), analogous y- and z- components have a similar structure, each involving a correspondent k-component dependence that needs to be integrated to obtain the full G_{RI} expression. In the following we will focus the effort to this task, k-components are involved into the binomial term $(ik_x-2\nu\overline{PQ}_x)$ of Eq.41, by expanding this as

$$(ik_x-2\nu\overline{PQ}_x)^{l_P+t-2\alpha} = \sum_{\beta=0}^{l_P+t-2\alpha}\frac{(l_P+t-2\alpha)!}{\beta!\,(l_P+t-2\alpha-\beta)!}(-1)^{l_P+t-\beta}(2\nu\overline{PQ}_x)^{l_p+t-2\alpha-\beta}(ik_x)^\beta \tag{42}$$

This expansion makes $G_{RI}^{(x)}$ more complex, but readable for a computational implementation. By defining $\xi \equiv (\gamma_P+\gamma_Q)/4\delta$ and splitting the exponential term in 41 as $\exp\big(\xi(ik_x-2\nu\overline{PQ}_x)^2\big) = \exp\big(-\xi(k_x^2+4i\nu\overline{PQ}_xk_x)\big)\exp\Big(\xi(4\nu^2\overline{PQ}_x^2)\Big)$. We get

$$\begin{aligned} G_{RI}^{(x)} = &\frac{\pi}{\sqrt{\delta}}\exp\Big(\xi(4\nu^2\overline{PQ}_x^2)\Big)\sum_{r=0}^{\left[\frac{l_Q}{2}\right]}\sum_{s=0}^{l_Q-2r}\sum_{t=0}^{s}\sum_{\alpha=0}^{\left[\frac{l_P+t}{2}\right]}\sum_{\beta=0}^{l_P+t-2\alpha} \\ &\frac{l_Q!\,(l_P+t)!}{r!(s-t)!\,t!\,\alpha!\,\beta!}\,\frac{-1^{l_P+l_Q+t-s-\beta}}{(l_P+t-2\alpha-\beta)!}\,(2\nu)^{s+l_P+t-2\alpha-\beta} \\ &\overline{PQ}_x^{s+l_P-2\alpha-\beta}\,(i)^{l_Q-2r-s+\beta}\,(k_x)^{l_Q-2r-s+\beta} \\ &\frac{(\eta)^{\alpha-l_Q+r}(\gamma_Q)^{l_P+t-2\alpha}}{2^{l_P+l_Q+t}\delta^{-P+t-\alpha}}\exp\big(-\xi(k_x^2+4i\nu\overline{PQ}_xk_x)\big) \end{aligned} \tag{43}$$

Here is worth to mention notation conventions, if the $l,\ m,\ n$ subscripts are respectively related to the x-, y-, z-components, $G_{RI}^{(y)}$ and $G_{RI}^{(z)}$ will have a similar form as Eq.43, then they will possess the respective labels $(r_l, s_l, t_l, \ldots, r_m, s_m, t_n, \ldots, r_n, s_n, t_n)$, etc (see Eq.46). If we denote $G_{RI}^{(x)'}$, $G_{RI}^{(x)'}$, $G_{RI}^{(x)'}$ as those given by 43 without the common factors, G_{RI} from equation 28 becomes

$$G_{RI} = \frac{1}{2\pi^2}\exp\Big(-\nu\vec{\boldsymbol{PQ}}^2\Big)\,\frac{\pi^3}{\delta^{3/2}}\,\exp\Big(\xi(4\nu^2\overline{PQ}^2)\Big)\int d\vec{\boldsymbol{k}}\frac{1}{\vec{\boldsymbol{k}}^2}G_{RI}^{(x)'}G_{RI}^{(y)'}G_{RI}^{(z)'} \tag{44}$$

The final expression for G_{RI} involves the integration in k-space, we can split this integral into its Cartesian terms, each with the same form. To achieve this, be the C_{RI}^{x} the k_x independent collection of terms in $G_{RI}^{(x)}$ given as:

$$C_{RI}^{x} = \sum_{r=0}^{\left[\frac{l_Q}{2}\right]}\sum_{s=0}^{l_Q-2r}\sum_{t=0}^{s}\sum_{\alpha=0}^{\left[\frac{l_P+t}{2}\right]}\sum_{\beta=0}^{l_P+t-2\alpha}(-1)^{l_P+l_Q+t-s-\beta}(i)^{l_Q-2r-s+\beta}(2\nu)^{s+l_P+t-2\alpha-\beta}(\overline{PQ}_x)^{s+l_P-2\alpha-\beta}$$
$$\left[\frac{l_Q!\,(l_P+t)!}{r!\,(s-t)!\,t!\,\alpha!\,\beta!\,(l_P+t-2\alpha-\beta)!}\right]\left[\frac{\eta^{\alpha-l_Q+r}\,\gamma_Q^{l_P+t-2\alpha}}{2^{l_P+l_Q+t}\delta^{l_P+t-\alpha}}\right] \tag{45}$$

then, a relatively simple structure follows,

$$G_{RI} = \frac{1}{2\pi^2}\,\frac{\pi^3}{\delta^{3/2}}\,\exp\Big((4\nu^2\xi-\nu)\vec{\boldsymbol{PQ}}^2\Big)\,C_{RI}^{x}C_{RI}^{y}C_{RI}^{z}\int\frac{d\vec{\boldsymbol{k}}}{\vec{\boldsymbol{k}}^2}\exp\Big(i\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{PQ}}\Big)$$
$$\exp\Big(\xi\vec{\boldsymbol{k}}^2\Big)\exp\Big(\xi 4i\nu\vec{\boldsymbol{PQ}}\cdot\vec{\boldsymbol{k}}\Big)\,k_x^{l_Q-2r_l-s_l+\beta_l}\,k_y^{m_Q-2r_m-s_m+\beta_m}\,k_z^{n_Q-2r_n-s_n+\beta_n} \tag{46}$$

recalling the identity: $\exp\Big(\epsilon\vec{\boldsymbol{k}}^2\Big) = 2\epsilon\vec{\boldsymbol{k}}^2\int_0^1 du\,u^{-3}\exp\Big(-\epsilon\vec{\boldsymbol{k}}^2u^{-2}\Big)$, and substitution into 46 leads to

$$G_{RI} = \frac{1}{\pi^2}\,\xi\,\frac{\pi^3}{\delta^{3/2}}\,\exp\Big((4\nu^2\xi-\nu)\vec{\boldsymbol{PQ}}^2\Big)\,C_{RI}$$
$$\int_0^1\int\frac{d\vec{\boldsymbol{k}}\,du}{u^3}\exp\Big(i(1-4\nu\xi)\vec{\boldsymbol{k}}\cdot\vec{\boldsymbol{PQ}}\Big)\exp\Big(-\xi\vec{\boldsymbol{k}}^2/u^2\Big)\,k_x^{\omega}\,k_y^{\theta}\,k_z^{\phi} \tag{47}$$

where $\omega \equiv l_Q - 2r_l - s_l + \beta_l$, $\theta \equiv m_Q - 2r_m - s_m + \beta_m$, $\phi \equiv n_Q - 2r_n - s_n + \beta_n$ are overall integral numbers computed as collections of indexes found in the sums, and $C_{RI} = C_{RI}^{x}C_{RI}^{y}C_{RI}^{z}$. The integral in k-space can be decomposed into components of the form

$$I_{k_x} = \int dk_x\exp\big(i(1-4\nu\xi)k_x\overline{PQ}_x\big)k_x^{\omega}\exp\big(-\xi k_x^2/u^2\big) \tag{48}$$

The solution of this type of integral is simple and follow a similar reasoning as those from Eq.31 as

$$I_{k_x} = \sqrt{\frac{\pi}{\xi}} \exp\left(-u^2\left[\frac{(1-4\nu\xi)^2\overline{PQ}_x^2}{4\xi}\right]\right) u \sum_{l=0}^{\omega} \frac{\omega!(-1)^l i^\omega}{i!(\omega-2i)!} \frac{\left[(1-4\nu\xi)\overline{PQ}_x\right]^{\omega-2l}}{2^\omega\, \xi^{\omega-l}} u^{2(\omega-l)} \tag{49}$$

The rest of the components I_{k_y} and I_{k_z} will resemble similar forms with respective m, n indexes in the sum and θ, ϕ sum limits. Be D_{RI}^x the collection of u-independent factors encountered in Eq.47:

$$D_{RI}^x = \sqrt{\frac{\pi}{\xi}} \sum_{l=0}^{\omega} \frac{\omega!(-1)^l i^\omega}{i!(\omega-2i)!} \frac{\left[(1-4\nu\xi)\overline{PQ}_x\right]^{\omega-2l}}{2^\omega\, \xi^{\omega-l}} \tag{50}$$

Grouping D_{RI}^x, D_{RI}^y and D_{RI}^z multiplicative factors into an overall D_{RI} we can finally write G_{RI} as:

$$\begin{aligned} G_{RI} = \frac{1}{\pi^2}\, \xi\, \frac{\pi^3}{\delta^{3/2}}\, \exp\Big((4\nu^2\xi - \nu)\vec{\boldsymbol{PQ}}^2\Big)\, C_{RI} D_{RI} \\ \int_0^1 \frac{du}{u^3} u^3 \exp\left(-\frac{(1-4\nu\xi)^2\overline{PQ}^2 u^2}{4\xi}\right) u^{2(\omega-l)} u^{2(\theta-m)} u^{2(\phi-n)} \end{aligned} \tag{51}$$

It is particularly convenient to pack the whole set of u-exponents into an overall exponent, $\Omega \equiv \omega - l + \theta - m + \phi - n$; and write G_{RI} as

$$\begin{aligned} G_{RI} = \frac{\xi}{\pi^2}\, \frac{\pi^3}{\delta^{3/2}}\, \exp\Big((4\nu^2\xi - \nu)\vec{\boldsymbol{PQ}}^2\Big)\, C_{RI} D_{RI} \\ \int_0^1 du\, u^{2\Omega} \exp\left(-\frac{(1-4\nu\xi)^2\overline{PQ}^2 u^2}{4\xi}\right) \end{aligned} \tag{52}$$

The use of Boys function

The exponential factor $-(1-4\nu\xi)^2/4\xi$ that appears in the integral of Eq.52 can be written in terms of γ_P and γ_Q, recalling our previous definitions; $4\xi = (\gamma_P+\gamma_Q)/\nu(\gamma_P+\gamma_Q)+\gamma_P\gamma_Q)$, then $(1-4\nu\xi)^2/4\xi = (\gamma_P^2\gamma_Q^2)/((\nu\gamma_P + \nu\gamma_Q + \gamma_P\gamma_Q)(\gamma_P+\gamma_Q)) \equiv \Xi$. The inner integral over u is the Boys function defined as

$$F_\Omega(\Xi\overline{PQ}^2) = \int_0^1 du\, u^{2\Omega} \exp\Big(u^2\overline{PQ}^2\, \Xi\Big) \tag{53}$$

G_{RI} is therefore

$$G_{RI} = \frac{\xi}{\pi^2}\frac{\pi^3}{\delta^{3/2}}\exp\Big((4\nu^2\xi-\nu)\vec{\boldsymbol{PQ}}^2\Big)\,C_{RI}D_{RI}F_{\Omega}(\Xi\,\overline{PQ}^2) \tag{54}$$

the required integrals can be written in terms of the so-called Boys function F_{Ω} [102]. The evaluation of this function can be made trough numerical integration with high precision suitable for electronic structure calculations. In out case this function is evaluated after calculation of $\Xi\overline{PQ}^2$, involving the information contained in the basis set (exponents, coefficients) and nuclear coordinates at which the primitives are centered. There is a collection of common factors contained in 54 which are independent of any sum index or integration variable and constitute an overall multiplicative term for G_{RI} given by

$$\frac{\xi}{\pi^2}\frac{\pi^3}{\delta^{3/2}}\exp\Big((4\nu^2\xi-\nu)\overline{PQ}^2\Big) = \frac{2\pi^{5/2}}{\delta(\gamma_P+\gamma_Q)^{1/2}}\exp\Big(-\frac{\nu\gamma_P\gamma_Q}{\delta}\overline{PQ}^2\Big) \tag{55}$$

Inserting 55 into G_{RI} gives the solution for the Gaussian two-electron damped integrals. However further steps are needed in order to derive an overall expression suitable for computational implementations, this will be shown in the following section.

Complete form of the gauERIs

The multiplicative terms $C_{RI}D'_{RI}$ can be arranged as

$$C^x_{RI}D'^x_{RI} = \sum_{r=0}^{\left[\frac{l_Q}{2}\right]} \sum_{s=0}^{l_Q-2r} \sum_{t=0}^{s} \sum_{\alpha=0}^{\left[\frac{l_P+t}{2}\right]} \sum_{\beta=0}^{l_P+t-2\alpha} (-1)^{l_P+l_Q+t-s-\beta} (i)^{l_Q-2r-s+\beta} (2\nu)^{s+l_P+t-2\alpha-\beta} (\overline{PQ}_x)^{s+l_P-2\alpha-\beta} \left[\frac{l_Q!\,(l_P+t)!}{r!\,(s-t)!\,t!\,\alpha!\,\beta!\,(l_P+t-2\alpha-\beta)!}\right] \left[\frac{\eta^{\alpha-l_Q+r}\,\gamma_Q^{l_P+t-2\alpha}}{2^{l_P+l_Q+t}\delta^{l_P+t-\alpha}}\right] \sum_{l=0}^{[\omega/2]} \frac{\omega!(-1)^l i^\omega}{i!(\omega-2i)!} \frac{\left[(1-4\nu\xi)\overline{PQ}_x\right]^{\omega-2l}}{2^\omega\,\xi^{\omega-l}} \tag{56}$$

for $\omega \equiv l_Q - 2r - s + \beta$ and rearranging terms we have

$$C^x_{RI}D'^x_{RI} = \sum_{r=0}^{\left[\frac{l_Q}{2}\right]} \sum_{s=0}^{l_Q-2r} \sum_{t=0}^{s} \sum_{\alpha=0}^{\left[\frac{l_P+t}{2}\right]} \sum_{\beta=0}^{l_P+t-2\alpha} \sum_{l=0}^{[\omega/2]} (2\nu)^{s+l_P+t-2\alpha-\beta} (1-4\nu)^{l_Q-2r-s+\beta} (\overline{PQ}_x)^{l_P+l_Q+s-2(\alpha+r+l)} \eta^{\alpha-l_Q+r} \gamma_Q^{l_P+t-2\alpha} \left[\frac{\delta^{l_Q+\alpha+\beta}}{\delta^{l_P+t+2r+s+l}}\right] \left[\frac{2^\beta}{2^{l_P+s+t+2r+2l}}\right] \left[\frac{1}{(\gamma_P+\gamma_Q)^{\omega-l}}\right] \left[\frac{l_Q!\,(l_P+t)!\,(-1)^{l_P+t+l}}{r!\,(s-t)!\,t!\,\alpha!\,\beta!\,(l_P+t-2\alpha-\beta)!}\right] \left[\frac{(l_Q-2r-s+\beta)!}{l!(l_Q-2r-s+\beta-2l)!}\right] \tag{57}$$

We can group these products with the corresponding Cartesian angular momentum coefficients C_{lP}, C_{lQ}, etc. (Eqs. 11 - 15) which are factors of G_{RI}. The overall product can be denoted as $C^{x*}_{RI}D^{x*}_{RI} = (C_{lP}C_{lQ}) \cdot (C^x_{RI}D^x_{RI})$, then G acquires the following final form

$$G = K_{AB}K_{CD}\frac{2\pi^{5/2}}{\delta(\gamma_P+\gamma_Q)^{1/2}} \exp\Big(-\frac{\nu\gamma_P\gamma_Q}{\delta}\overline{PQ}^2\Big) \cdot \sum_{l_P}^{l_A+l_B} \sum_{l_Q}^{l_C+l_D} C^{x*}_{RI}D^{x*}_{RI} \sum_{m_P}^{m_A+m_B} \sum_{m_Q}^{m_C+m_D} C^{y*}_{RI}D^{y*}_{RI} \sum_{n_P}^{n_A+n_B} \sum_{n_Q}^{n_C+n_D} C^{z*}_{RI}D^{z*}_{RI}F_\Omega(\Xi\,\overline{PQ}^2). \tag{58}$$

Fundamental gauERI

The fundamental Gaussian damped electron repulsion integral denoted as $\langle ss| V_{eff} |ss\rangle$ corresponds to the case in which the GTOs are of zero angular momentum. Recalling equation 55, the integral reduces to

$$\begin{aligned} \langle ss| V_{eff} |ss\rangle = \frac{2\pi^{5/2}}{(\nu\gamma_P + \nu\gamma_Q + \gamma_P\gamma_Q)(\gamma_P + \gamma_Q)^{1/2}} \\ \cdot \exp\left(-\frac{\gamma_P\gamma_Q\nu\overline{PQ}^2}{\nu\gamma_P + \nu\gamma_Q + \gamma_P\gamma_Q}\right) \cdot F(\frac{\overline{PQ}^2\gamma_P^2\gamma_Q^2}{(\nu\gamma_P + \nu\gamma_Q + \gamma_P\gamma_Q)(\gamma_P + \gamma_Q)}) \end{aligned} \tag{59}$$

This is the same expression as that derived by Boys [103] (theorem 8, page 3).

Additional information for atomic systems

Correlation energies computed for atoms with different μ ν values

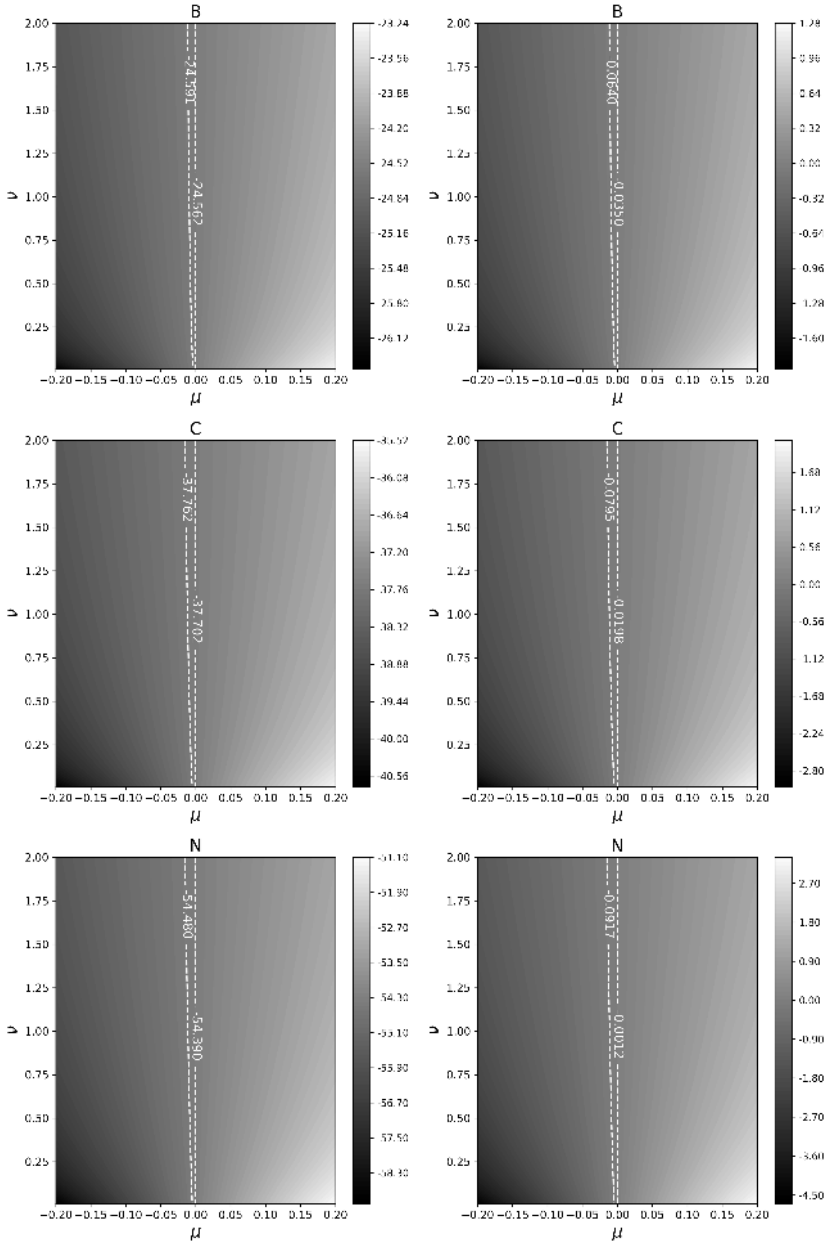


Figure 25: B, C and N atoms total (left plots) and correlation (right plots) energies (in a.u.) computed within the CASSCF+V_{eff} represented as contour plots resulting from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot.

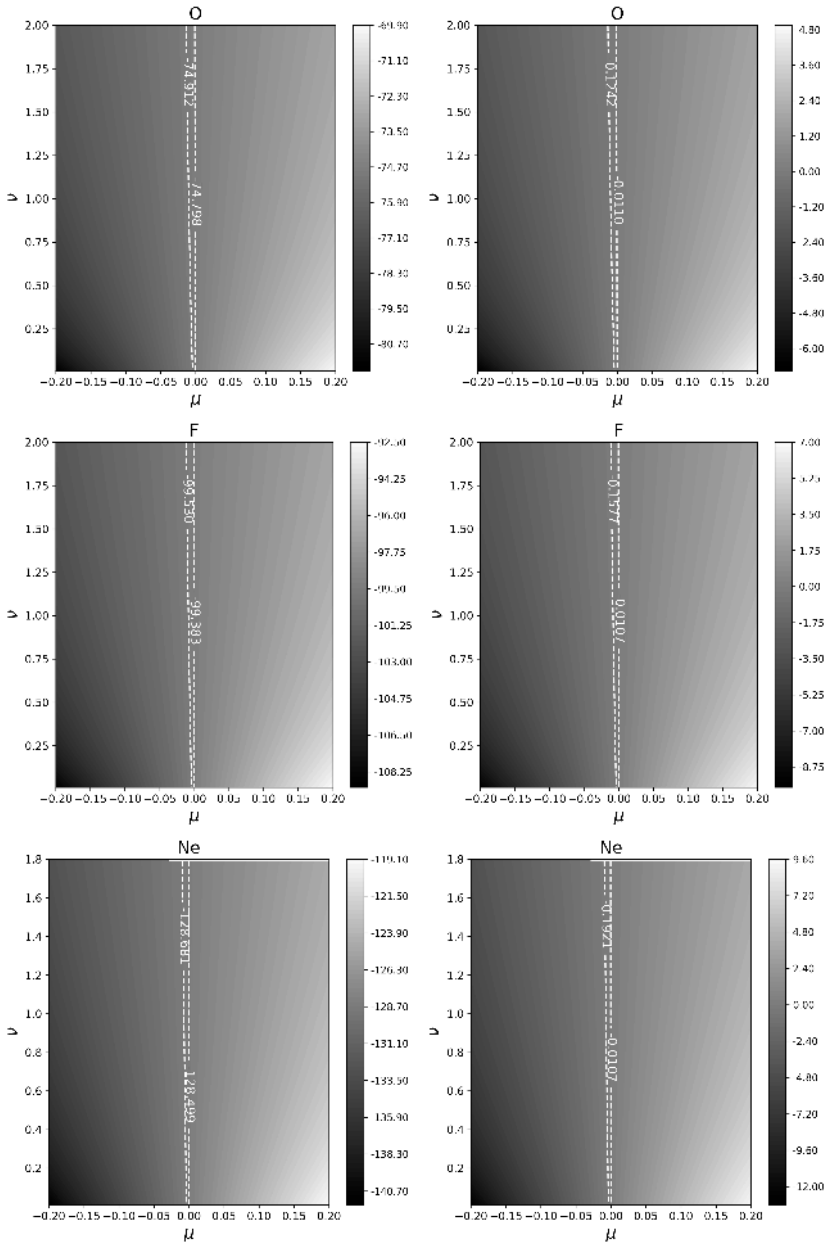


Figure 26: O, F and Ne total (left plots) and correlation (right plots) energies (in a.u.) computed within the CASSCF+V_{eff} represented as contour plots resulting from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot.

Total CASSCF+V_{eff} and corresponding correlation energies computed for cations with different μ ν values

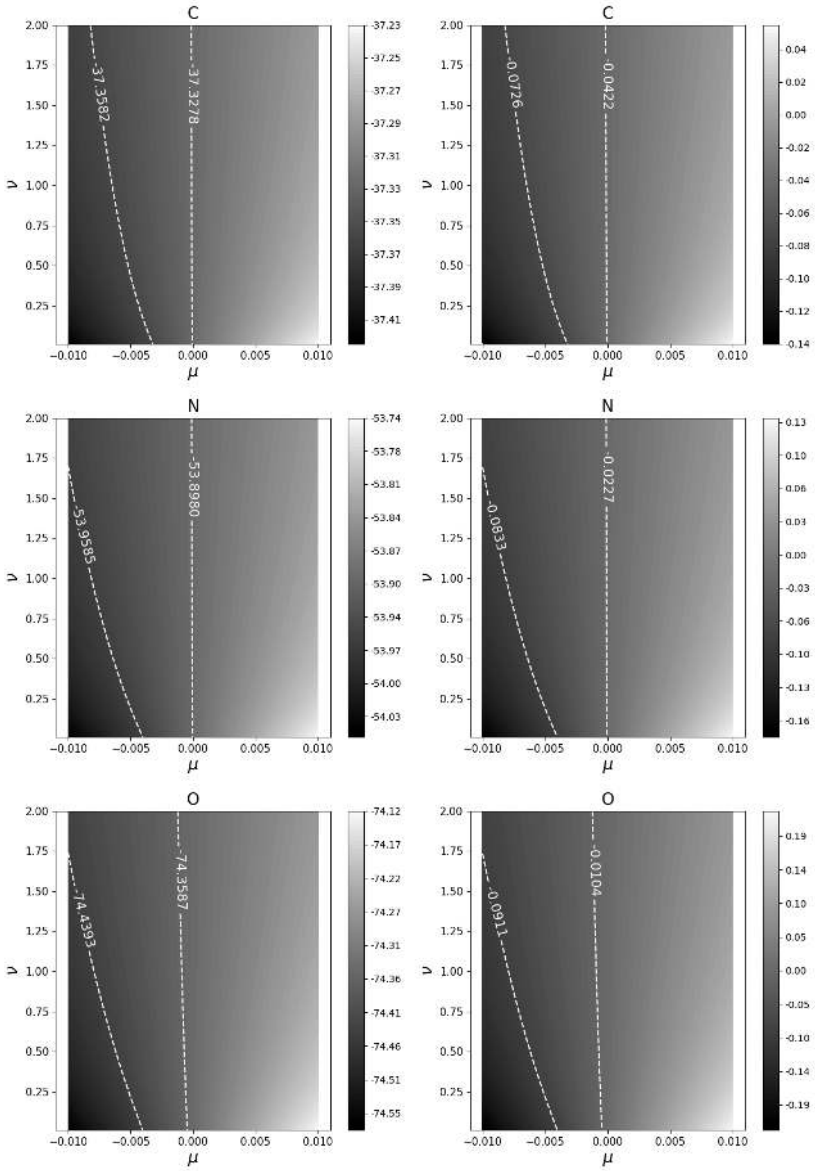


Figure 27: C^+, N^+, O^+ cations CASSCF+V_{eff} obtained contour plots of total (left plots) and correlation energies (right plots), from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot.

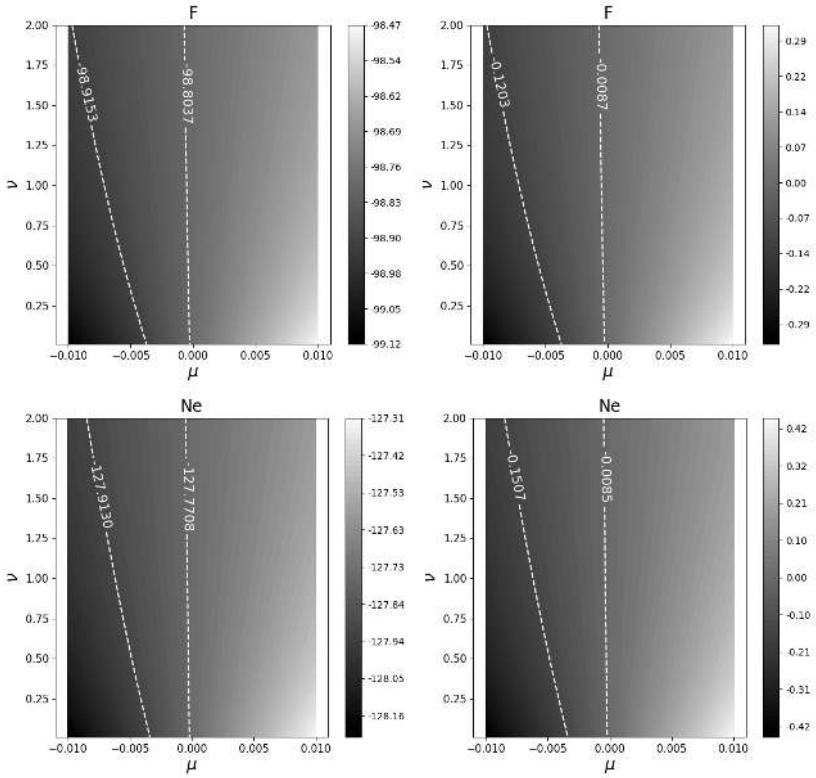


Figure 28: F^+, Ne^+, cations CASSCF+V_{eff} obtained contour plots of total (left plots) and correlation energies (right plots), from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot.

Total CASSCF+V_{eff} and corresponding correlation energies computed for anions with different μ ν values

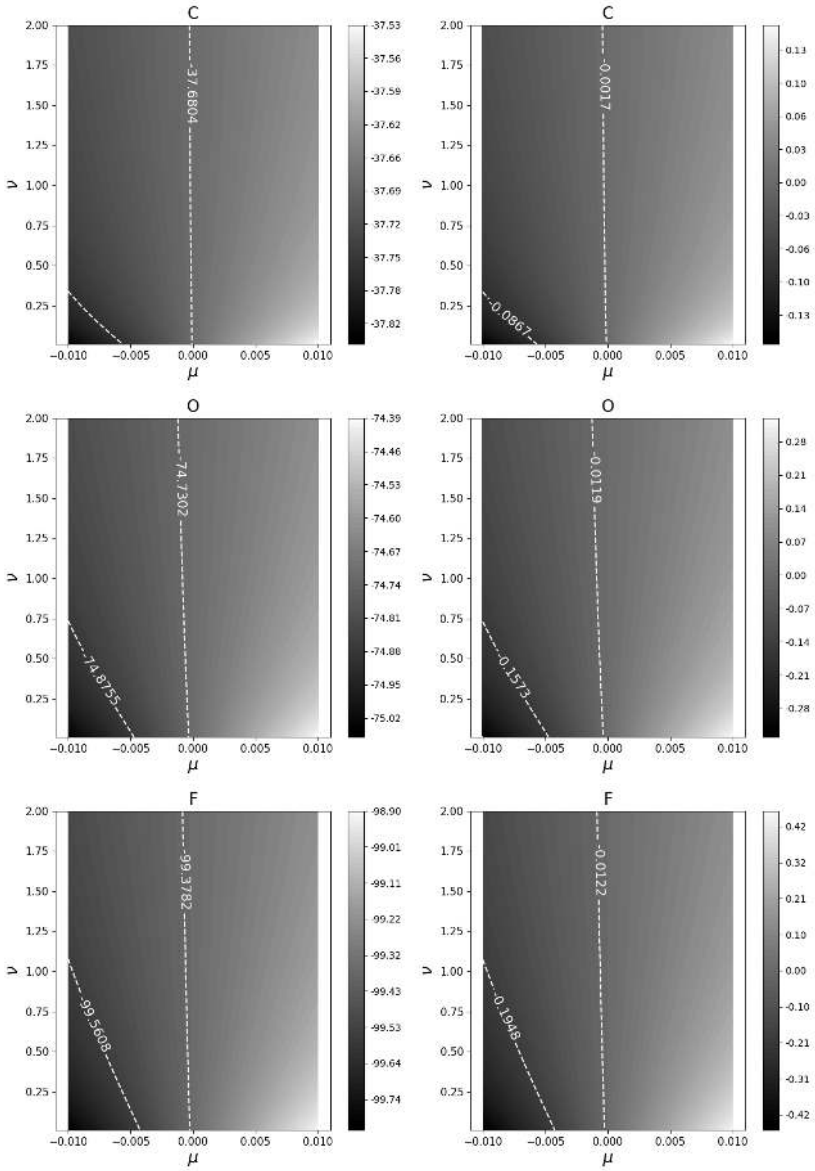


Figure 29: Anions CASSCF+V_{eff} obtained contour plots of total (left plots) and correlation energies (right plots), from variations of μ and ν parameters, FCI and pristine CASSCF energies are indicated in green dashed lines in each plot.

Table 25: Optimal parameters for atomic systems computed from a linear fitting from computed data (variations of μ and ν parameters).

Atom	ν	μ	E_c (a.u.)
B	0.2623	-0.006	-0.0632
C	0.1241	-0.006	-0.0786
N	0.1632	-0.006	-0.0909
O	0.2567	-0.006	-0.1233
F	0.4157	-0.006	-0.1572
Ne	0.6848	-0.006	-0.1886
C+	0.9502	-0.006	-0.0708
N+	0.4540	-0.006	-0.0811
O+	0.4853	-0.006	-0.0891
F+	0.6803	-0.006	-0.1176
Ne+	0.9678	-0.006	-0.1479
C-	0.0340	-0.006	-0.0861
O-	0.1731	-0.006	-0.1560
F-	0.3131	-0.006	-0.1934

Coulomb holes for C atom

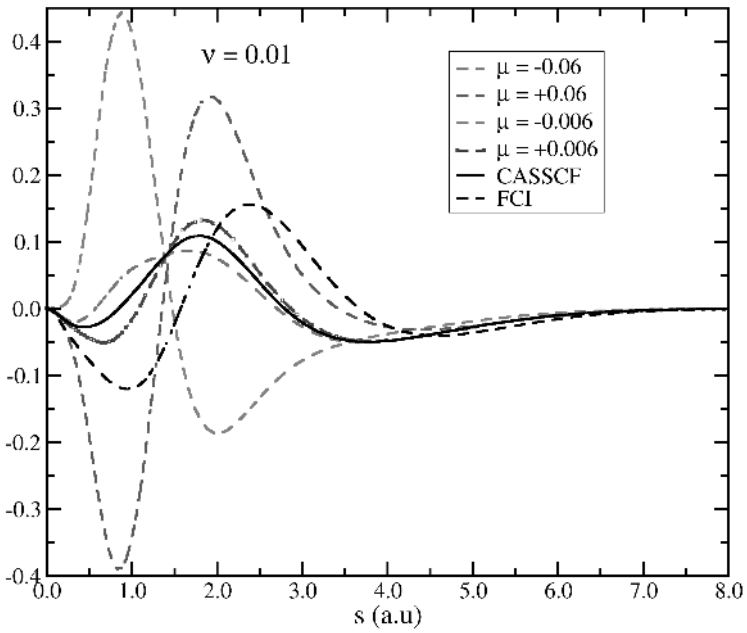


Figure 30: Coulomb holes for the C atom with fixed $\nu = 0.01$ value and variations of μ (both positive and negative), computed with the cc-pVDZ basis set.

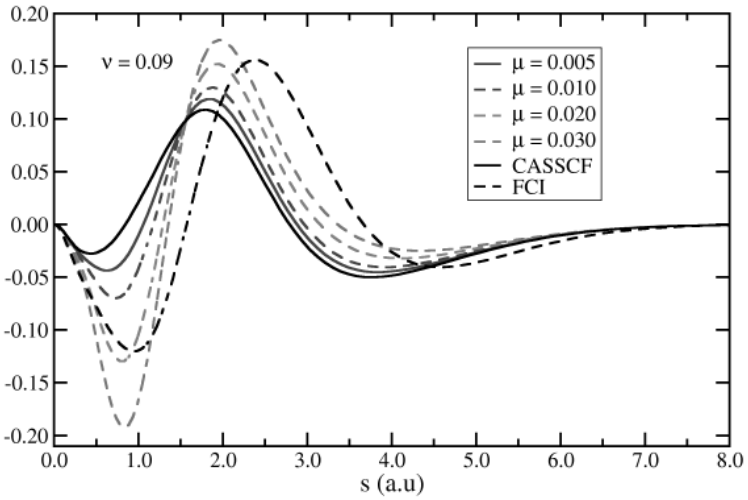


Figure 31: Evaluation of the effect of a positive effective potential. Coulomb holes for the C atom with fixed $\nu = 0.09$ value and variations of μ, computed with the cc-pVDZ basis set.

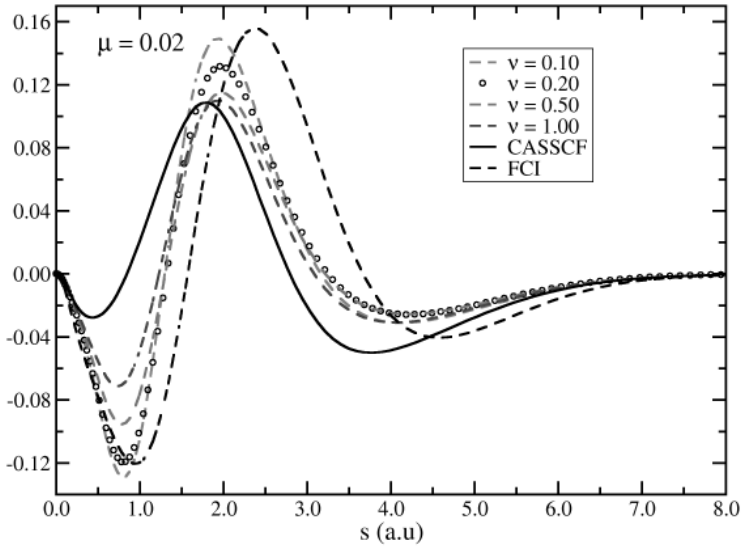


Figure 32: Evaluation of the effect of a positive effective potential. Coulomb holes for the C atom with fixed $\mu = 0.02$ value and variations of ν.

Additional information for diatomic systems

CASSCF+$V_{\mathbf{eff}}$ PES and Coulomb holes for H_2

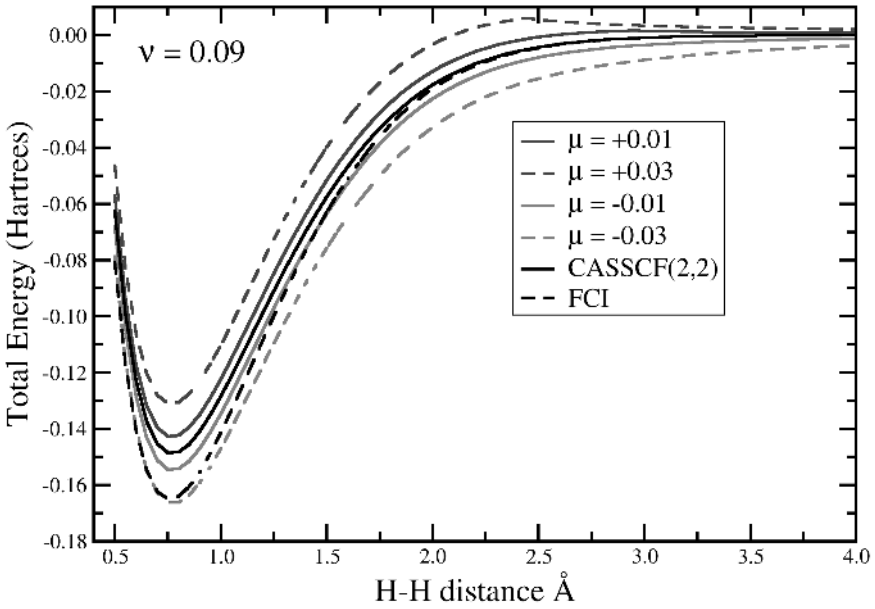


Figure 33: H_2 PECs computed with different values of μ (positive and negative), small variations lead to considerable changes in the PES, $\nu = 0.09$. Here the use of $\mu = +0.03$ leads to apparent distortions at long-range regions in the PEC (dashed blue line).

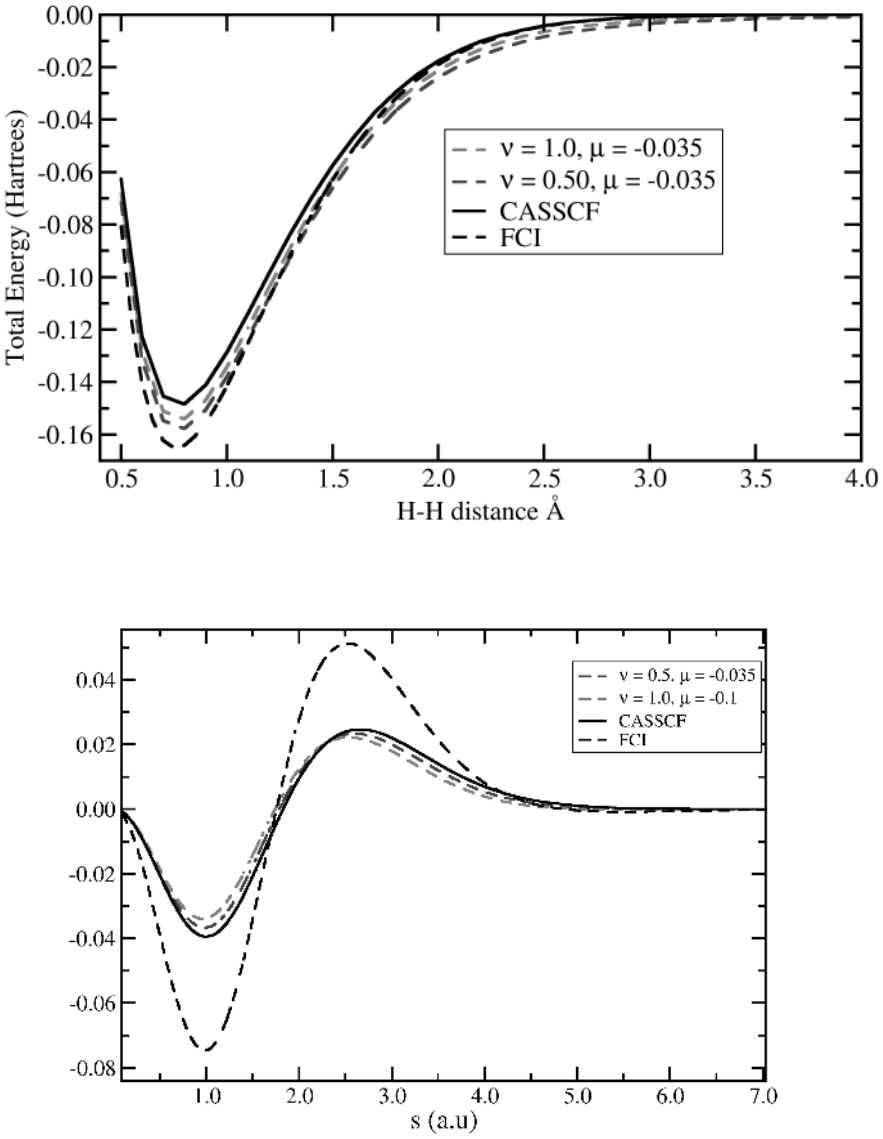


Figure 34: (Up) Potential energy surfaces of H_2 computed with selected parameters yielding energles close to FCI/cc-pVDZ. Changes in ν do not lead to large variations of the PES. (Down) Corresponding Coulomb holes at equilibrium. Small varitions do not deteriorate the Coulomb hole with respect to CASSCF.

Results for Isoelectronic Diatomics: N_2 and NO^+

The following PES were obtained withihn CASSCF(6,6)+V_{eff}/cc-pVDZ calculations. The effective potential parameters were optimized to yield the lowest error in the dissociation energies as indicated in main text.

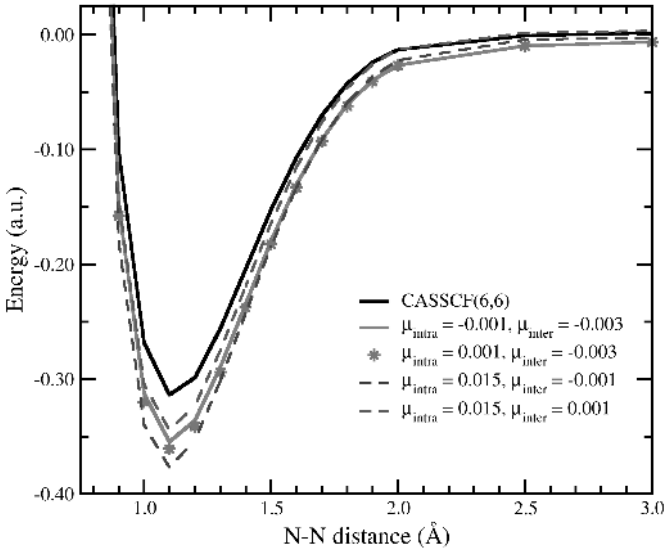


Figure 35: Best estimates computed for N_2 PECs, giving dissociation energies close to experimental values taken from Ref. [2] (calculations performed with $\nu = 0.5$ fixed value).

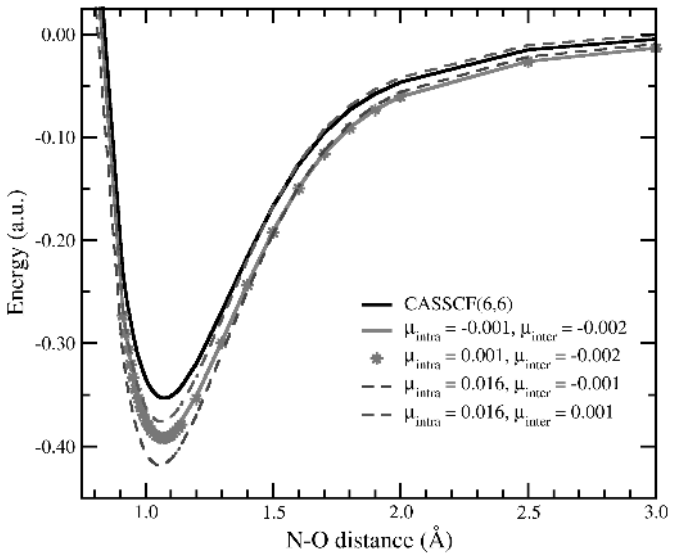


Figure 36: Best estimates computed for NO^+ PECs, giving dissociation energies close to experimental values taken from Ref. [2] (calculations performed with $\nu = 0.5$ fixed value).

Results for Triplet State Diatomics: O_2 and NH

For the triplet dimers examined in this study: O_2 and NH, we conducted CASSCF(4,4)/cc-pVDZ calculations. Computed energies exhibit less variation in response to changes in

the ν parameter compared to singlets, resulting in similar potential energy surfaces (PES) for different ν values. For these systems, the CASSCF errors in dissociation energy, when compared to experimental values (from Ref. [2]), are approximately 25-28%. However, by computing dissociation energies within the effective potential, this error is substantially reduced (see Figs.37 and 38). Up to date, Coulomb holes are not available for these systems (and also for the double OH case).

Best estimates for O_2 dissociation.

An active space with 8 electrons in six orbitals was used for this calculation. The following plot concerns to the best PES obtained within the CASSCF+V_{eff} method. The dissociation energy was taken as the zero of energy for each case.

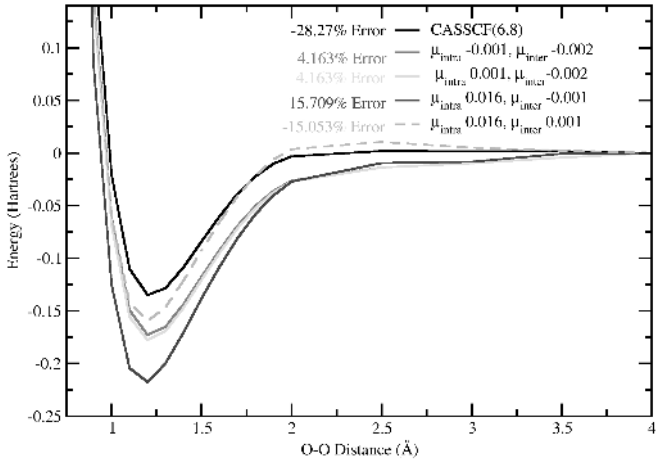


Figure 37: The lowest value considered $\nu = 0.1$ gives the best PES for O_2. Percentage of errors with respect to experimental dissociation values (taken from Ref. [2])follows the same color code as that indicated by labels in the plot.

Best estimates for NH dissociation.

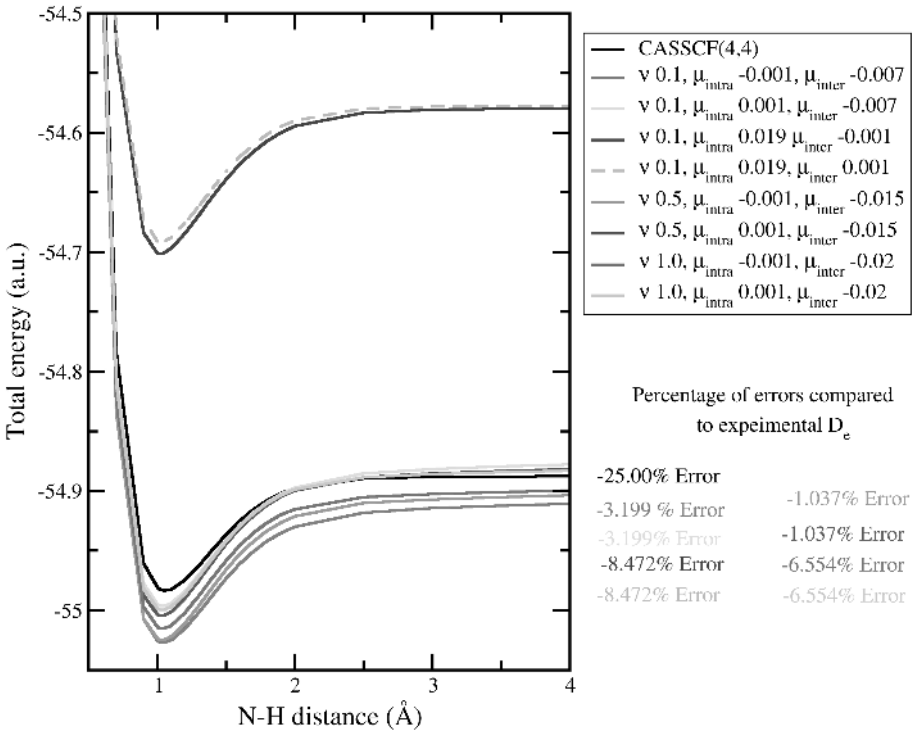


Figure 38: Selected PES for NH case, labels show the values of the two μ parameters in the effective potential together with the ν selection. Percentage of errors with respect to experimental dissociation values (from Ref. [2]) follows the same color code as that indicated by labels in the plot.

Results for Doublet State Diatomics: OH

Additionally, we explore the Potential Energy Curves (PECs)for the doublet OH dissociation, based on CASSCF(4o,5e) calculations. CASSCF underestimates dissociation energy with errors of the order of 20.62% as in the previous case. The inclusion of the effective potential significantly reduces this error to orders of -1.04%. In all of our most accurate PECs estimates, the inter-atomic contribution to the effective potential is larger (and negative) compared to the intra-atomic contribution.

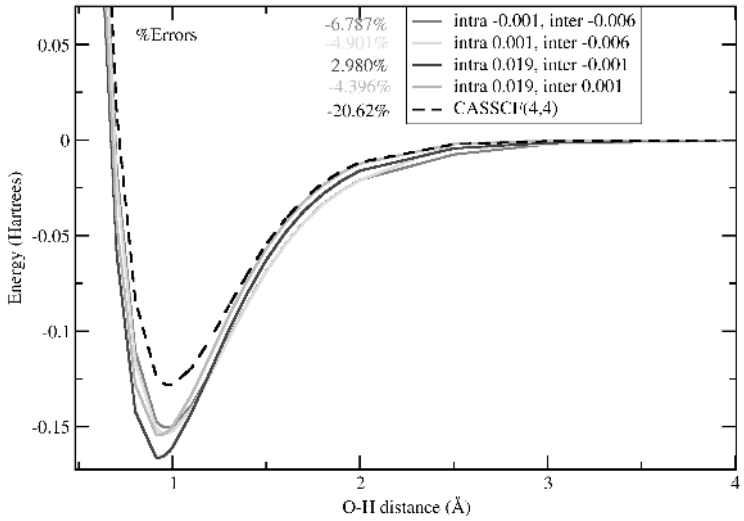


Figure 39: Selected PES for OH doublet, results computed with $\nu = 0.1$. As in previous examples, the dissociation errors are depicted withing the same color code as the plot.

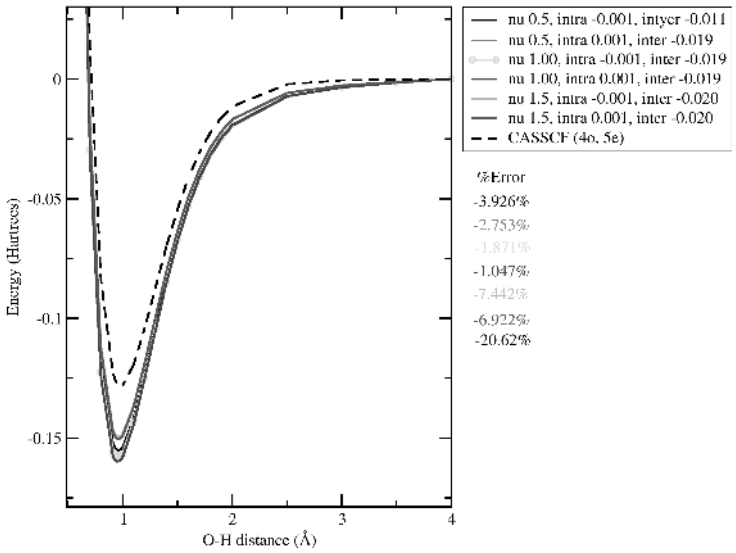


Figure 40: Selected PES for OH doublet, results computed with different values of ν exponential parameter, small changes to the PES are obtained for the different combinations of the multiplicative μ parameters. As in previous examples, the dissociation errors are depicted withing the same color code as the plot.

Additional information regarding parametrization strategy 2 (PS2) and data fitting.

Here, recalling erf-fitting from PS2 results, we compute additional QTAIM data in order to try to stablish a relation between system's QTAIM quantities (delocalization index (D(A,B)) and atomic volumes) and the optimal μ_{inter} patterns. For equilibrium geometries, corresponding a, b and R^2 values are depicted as bar plots in Figure 41, where we include obtained These have been computed in order to elucidate some trend related with erf-fitting parameters (a, b). Qualitatively, in most cases, b parameter follows a similar tendency as D(A,B), while

a displays small values that show small variations with respect to changes in both ν and μ_{intra} values. There is a tendency for the b parameter to decrease as μ_{intra} increases. This effect is not general but becomes larger when there is a significant difference in the volume of the atoms as in the NH molecule. Corresponding numerical data can be consulted in Table 26

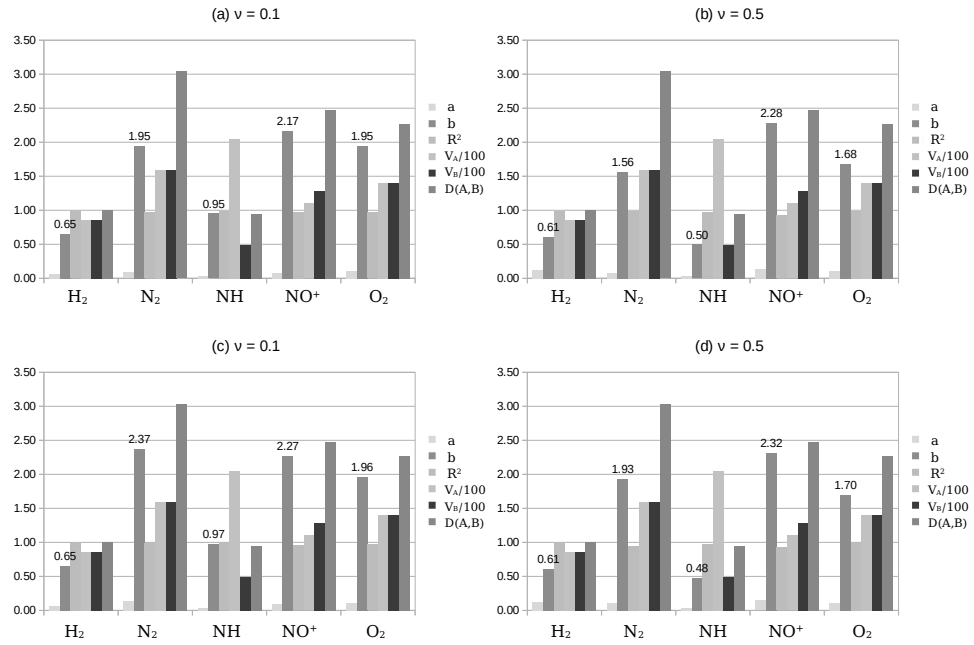


Figure 41: Fitted parameters and comparison with a fraction of the atomic volumes ($V_A/100$ and $V_B/100$) and delocalization index (D(A,B)) computed at equilibrium. Results computed with different ν values as indicated, $\mu_{intra} = 0.000$ (for (a) and (b) cases) and $\mu_{intra} = 0.001$ (for (c) and (d) cases).

Table 26: Full table with optimal fitting values to the function $a(erf(bR_{AB}-1))$, R^2 value and their comparison with atomic volumes (V_A and V_B) and delocalization index DI(A,B) obtained as indicated in main text.

Molecule	μ_{intra}	ν	a	b	R^2	V_A/100	V_B/100	DI(A,B)
H2		0.1	0.0610	0.6493	0.9953	0.8589	0.8589	1.0000
		0.1	0.0610	0.6493	0.9953	0.8589	0.8589	1.0000
		0.1	0.0600	0.6093	0.9929	0.8589	0.8589	1.0000
		0.5	0.1217	0.6115	0.9941	0.8589	0.8589	1.0000
		0.5	0.1217	0.6115	0.9941	0.8589	0.8589	1.0000
		0.5	0.1218	0.6095	0.9939	0.8589	0.8589	1.0000
N2	0.000	0.1	0.0834	1.9469	0.9649	1.5933	1.5933	3.0393
	0.001	0.1	0.1328	2.3728	0.9900	1.5933	1.5933	3.0393
	0.010	0.1	0.1352	2.3796	0.9848	1.5933	1.5933	3.0393
	0.000	0.5	0.0776	1.5614	0.9906	1.5933	1.5933	3.0393
	0.001	0.5	0.1123	1.9267	0.9469	1.5933	1.5933	3.0393
	0.010	0.5	0.1246	1.9709	0.9164	1.5933	1.5933	3.0393
NH	0.000	0.1	0.0301	0.9542	0.9849	2.0498	0.4921	0.9444
	0.001	0.1	0.0294	0.9735	0.9834	2.0498	0.4921	0.9444
	0.010	0.1	0.0217	1.1404	0.9680	2.0498	0.4921	0.9444
	0.000	0.5	0.0343	0.4959	0.9762	2.0498	0.4921	0.9444
	0.001	0.5	0.0334	0.4811	0.9775	2.0498	0.4921	0.9444
	0.010	0.5	0.0410	0.9410	0.9536	2.0498	0.4921	0.9444
NO+	0.000	0.1	0.0788	2.169	0.9709	1.1006	1.2801	2.47261
	0.001	0.1	0.08987	2.267	0.9544	1.1006	1.2801	2.47261
	0.010	0.1	0.2842	3.0964	0.9668	1.1006	1.2801	2.47261
	0.000	0.5	0.1376	2.2844	0.9255	1.1006	1.2801	2.47261
	0.001	0.5	0.1454	2.3203	0.9241	1.1006	1.2801	2.47261
	0.010	0.5	0.1194	2.1375	0.9671	1.1006	1.2801	2.47261
O2	0.000	0.1	0.1043	1.9469	0.9649	1.4056	1.4056	2.26941
	0.001	0.1	0.1052	1.9627	0.9678	1.4056	1.4056	2.26941
	0.010	0.1	0.1011	2.0515	0.985	1.4056	1.4056	2.26941
	0.000	0.5	0.1041	1.6801	0.9901	1.4056	1.4056	2.26941
	0.001	0.5	0.1063	1.6968	0.9905	1.4056	1.4056	2.26941
	0.010	0.5	0.1188	1.8396	0.9968	1.4056	1.4056	2.26941

Appendix D

Mean-Field Spin-Orbit Coupling

In the mean-field approximation, the two-electron part of the BP Hamiltonian can be written as:

$$\begin{aligned} H^{SOMF} &= \sum_{pq} \left[I_{pq} + \sum_{rs} \rho_{rs} (J_{prqs} - J_{prsq} - J_{rpqs}) \right] a_p^\dagger a_q \\ &= \sum_{pq} \left[h_{pq}^{1eSO} + h_{pq}^{2eSOMF} \right] a_p^\dagger a_q \end{aligned} \tag{60}$$

where h_{pq}^{1eSO} and h_{pq}^{2eSOMF} are the spin-orbit one and two-electron mean-field matrices, ρ_{rs} is a density matrix, and I_{pq} and J_{prqs} respectively are one- and two-electron spin-orbit integrals,

$$I_{pq} = \langle \phi_p | \mathbf{h}^{SO}(1) \cdot \mathbf{s}(1) | \phi_q \rangle \tag{61}$$

$$J_{pqrs} = \langle \phi_p(1) \phi_q(2) | \mathbf{h}^{SOO}(1,2) \cdot (\mathbf{s}(1) + 2\mathbf{s}(2)) | \phi_r(1) \phi_s(2) \rangle \tag{62}$$

By expanding contractions, and integrating over spin variables in equations (61) and (62), the mean field Hamiltonian can be written in terms of the sum of one- and two-electron components denoted as h_x^{SOMF}, h_y^{SOMF}, h_z^{SOMF} as:[456]

$$\begin{aligned} H^{SOMF} = \frac{1}{2} \sum_{pq} [& h_{L+,pq}^{SOMF} a_{\bar{p}}^\dagger a_q + h_{z,pq}^{SOMF} a_p^\dagger a_q \\ & + h_{z,\bar{p}\bar{q}}^{SOMF} a_{\bar{p}}^\dagger a_{\bar{q}} + h_{L-,p\bar{q}}^{SOMF} a_p^\dagger a_{\bar{q}}], \end{aligned} \tag{63}$$

where the we assume restricted spin orbitals. The creation and annhilation operators of α and β spin orbitals are indicated by indices without and with overbars respectively, the spherical components that appear in equation (63) are defined by

$$h_{L+,pq}^{SOMF} = h_x^{SOMF} + i h_y^{SOMF} \tag{64}$$

$$h_{L-,pq}^{SOMF} = h_x^{SOMF} - i h_y^{SOMF} \tag{65}$$

When the following condition is satisfied,

$$h_{z,pq}^{SOMF} a_p^\dagger a_q = -h_{z,\bar{p}\bar{q}}^{SOMF} a_{\bar{p}}^\dagger, \tag{66}$$

the Hamiltonian of equation (63) can be written in terms of triplet excitation operators as:

$$H^{SOMF} = \frac{1}{2} \sum_{pq} [h_{L+,pq}^{SOMF} T_{pq}^{1,-1} + \sqrt{2} h_{z,pq}^{SOMF} T_{pq}^{1,0} - h_{L-,p\bar{q}}^{SOMF} T_{pq}^{1,1}], \tag{67}$$

where $T_{pq}^{1,-1}$, $T_{pq}^{1,0}$ and $T_{pq}^{1,1}$ are defined as:

$$\begin{aligned} T_{pq}^{1,-1} &= a_{\bar{p}}^{\dagger} a_q \\ T_{pq}^{1,-1} &= \frac{1}{\sqrt{2}}(a_p^{\dagger} a_q - a_{\bar{p}}^{\dagger} a_{\bar{q}}) \\ T_{pq}^{1,-1} &= -a_p^{\dagger} a_{\bar{q}} \end{aligned} \tag{68}$$

The condition stated in equation (66) is satisfied by the one-electron terms in the Hamiltonian, but this is not always true for the two-electron mean-field part. Concretely, equation (66) is satisfied only for reference density matrices ρ with equal α and β contributions. Consequently, only the use of densities from spin-pure states with same number of α and β electrons is allowed in order to guarantee the symmetry of the operator. Non-physical contributions to the SOC are introduced with the use of different spin-projected densities ρ_{pq} in equation (60).

Comparative Results for Trimethylene

The different components of SOC and SOCCs for trimethylene are compiled in Tables 27 and 28, Table 29 contains the SOCCs computed with the 6-31G basis set for the triplet $M_s = 0$, and $M_s = 1$, labeled respectively as SOC(0) and SOC(1).

Table 27: SOC components and constants (in cm^{-1}) for the different geometries of TM. Calculations performed with different basis sets and a (6e,5o) RAS2 space.

	cc-pVDZ		cc-pVTZ		cc-pVQZ	
Geometry (0,0)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	0.0256	0.0000	-0.0297	0.0000	-0.0215	0.0000
Real 2e	-0.0321	0.0000	0.0350	0.0000	0.0252	0.0000
Imag. 1e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Imag. 2e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Total	0.0065	0.0000	0.0054	0.0000	0.0000	0.0000
Geometry (0,90)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	0.1152	0.0000	0.1146	0.0000	0.0811	0.0000
Real 2e	-0.0591	0.0000	-0.0571	0.0000	-0.0400	0.0000
Imag. 1e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Imag. 2e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Total	0.0794	0.0000	0.0575	0.0000	0.0411	0.0000
Geometry (90,90)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	-5.8375	0.0000	-6.1230	0.0000	-4.3974	0.0000
Real 2e	2.5934	0.0000	2.6670	0.0000	1.8974	0.0000
Imag. 1e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Imag. 2e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Total	3.2441	0.0000	3.4561	0.0000	2.5000	0.0000
Geometry (150,60)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	-0.9211	0.1364	-0.9886	0.1448	-1.1631	0.1492
Real 2e	0.3842	-0.0679	0.4040	-0.0714	0.4970	-0.0729
Imag. 1e	0.0000	0.7436	0.0000	0.8029	0.0000	0.7176
Imag. 2e	0.0000	-0.3276	0.0000	-0.3465	0.0000	-0.2904
Total	0.5369	0.4216	0.5846	0.4623	0.6661	0.4340

Table 28: Continuation, SOC components and constants (in cm^{-1}) for the different geometries of TM. Calculations performed with different basis sets and a (6e,5o) RAS2 space

	cc-pVDZ		cc-pVTZ		cc-pVQZ	
Geometry (145,35)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	-0.4426	0.4415	-0.4849	0.4912	-0.7210	0.3563
Real 2e	0.1735	-0.1961	0.1863	-0.2140	0.3112	-0.1356
Imag. 1e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Imag. 2e	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Total	0.3805	0.2454	0.4223	0.2773	0.4098	0.2207
Geometry (35,90)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	1.8280	0.8983	1.9309	0.3554	1.3611	0.3616
Real 2e	-0.8023	-0.3900	-0.8311	-0.1756	-0.5716	-0.1771
Imag. 1e	0.0000	0.3415	0.0000	0.9477	0.0000	1.3894
Imag. 2e	0.0000	-0.1708	0.0000	-0.4023	0.0000	-0.5924
Total	1.4505	0.5362	1.5554	0.5743	0.7895	0.8181
Geometry (45,60)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	-2.1300	-0.3122	-2.2238	-0.4592	-0.4638	-0.4637
Real 2e	0.9561	0.1340	0.9796	0.2258	0.1920	0.2259
Imag. 1e	0.0000	-0.4483	0.0000	-0.3244	0.0000	-0.1593
Imag. 2e	0.0000	0.2234	0.0000	0.1358	0.0000	0.6954
Total	1.6601	0.2869	1.7596	0.3001	0.2718	0.9283
Geometry (35,35)	SOC(0)	SOC(1)	SOC(0)	SOC(1)	SOC(0)	SOC(1)
Real 1e	-0.6399	-0.1809	-0.6646	-0.1775	-0.1751	0.0000
Real 2e	0.3085	0.0890	0.3168	0.0862	0.0844	0.0000
Imag. 1e	0.0000	0.0000	0.0000	0.0000	0.4740	0.0000
Imag. 2e	0.0000	0.0000	0.0000	0.0000	-0.2245	0.0000
Total	0.4687	0.1299	0.4919	0.0912	0.2654	0.0000

Table 29: Systematic comparison of SOCCs (in cm^{-1}), effect of the RAS2 active space indicated in parenthesis computed with the 6-31G basis set.

(α, β)	(2e,2o)	(4e,3o)	(6e,5o)
(0,0)	0.0028	0.0097	0.0070
(90,0)	0.0733	0.0662	0.0557
(90,90)	2.6817	2.7177	3.0927
(150,60)	0.6029	0.6335	0.7300
(145,35)	0.3000	0.3165	0.3784
(35,90)	1.0396	1.0869	1.2053
(45,60)	1.0427	1.0909	1.1965
(35,35)	0.3025	0.3222	0.3562

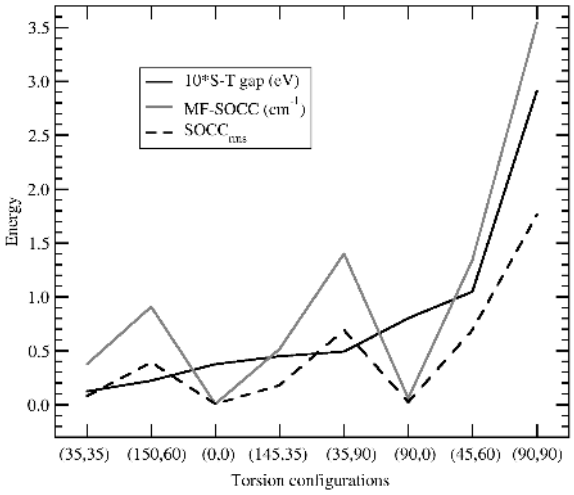
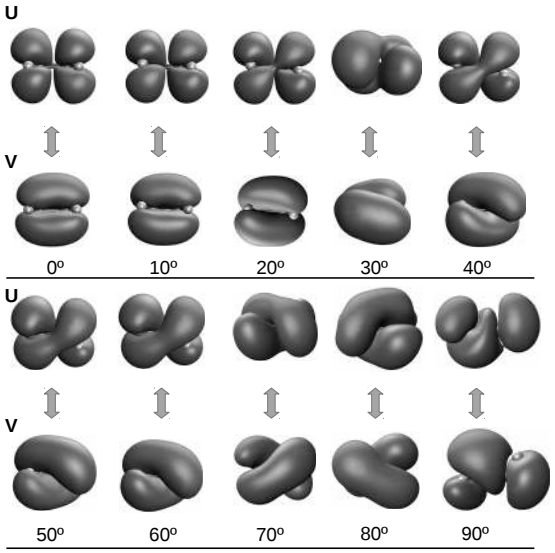


Figure 42: Comparison of results, with cc-pVQZ basis set, among the SOCC, the ST gap and the SOC_{rms} for TM. The energy scale is shown only for illustrative purposes, SOCCs are in cm^{-1} while the values of the vertical S-T gaps are presented as ten times their original value (in eV). Similar qualitative trends are found by the two approaches.

Complementary data for the C_2H_4 model

Natural transition orbitals for the C_2H_4 torsion model:



	ω_1	ω_2	PR
0º	0.9233	0.0482	1.112
10º	0.9229	0.0487	1.153
20º	0.9229	0.0489	1.113
30º	0.9178	0.0537	1.124
40º	0.9115	0.0600	1.390
50º	0.8998	0.0720	1.168
60º	0.8774	0.0957	1.225
70º	0.8292	0.1469	1.356
80º	0.7186	0.2627	1.662
90º	0.4927	0.4927	2.022

Figure 43: Dominant NTO pairs of the spinless transition matrix of ethylene (ROHF/cc-pVTZ triplet as reference) in the $T_1 \leftarrow S_0$ transitions at different torsion angles (isovalues of 0.05). The two largest singular values (ω_1 and ω_2) and participation ratios (PR) are presented on the right.

ROHF orbitals from the reference state. Evolution along the dissociation of the C_2H_4 model. The reduction of the two CH_2 separation induces a rearrangement in the order of the RAS2 reference orbitals as illustrated in Figs. 44 and 45.

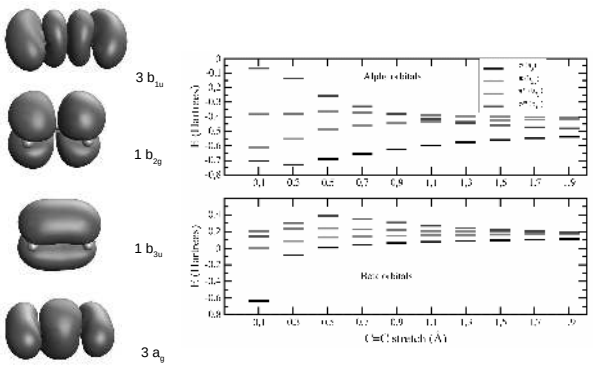


Figure 44: Reference ROHF quintet orbitals (which forms RAS2) for the D_{2d} configuration, and its evolution along the C=C stretching.

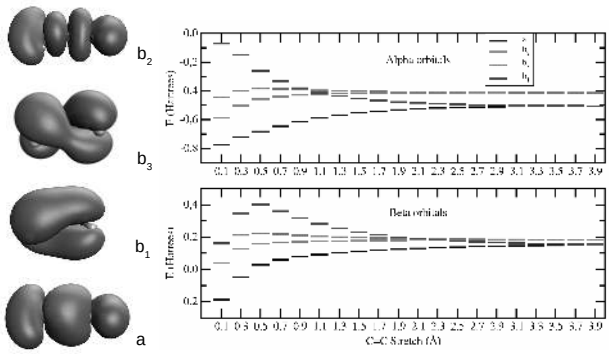


Figure 45: Reference ROHF quintet orbitals (which forms RAS2) for the D_2 configuration, and its evolution along the C=C stretching.

Dissociation curves for higher energy states. The dissociation curves for the twelve lowest energy states at dissociation, for the $0°$ twist-case are depicted in Figure 46, those for the $50°$ twist-case are depicted in Figure 47. In both figures, the blue-color curves indicate the states that are coupled with the three lowest energy states at dissociation. As shown, the singlet states ^1A_g and ^1A, remain as the lowest energy state along each respective dissociation curve, with the gap to the lowest triplet and quintet increasing as the C=C length is shortened.

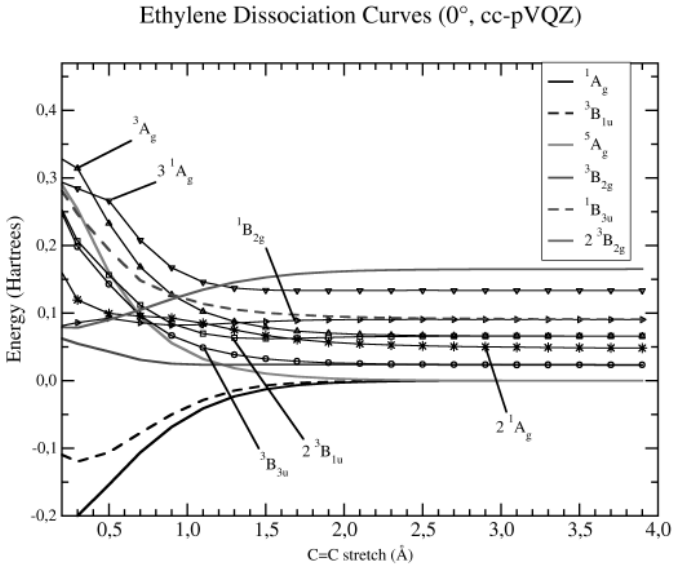


Figure 46: Dissociation curves for ethylene at $0°$ twist configuration computed with cc-pVQZ basis set. The graph shows the twelve selected states, those with lines in blue (and violet) are the ones that have non-zero spin orbit coupling with the three lowest lying states at dissociation

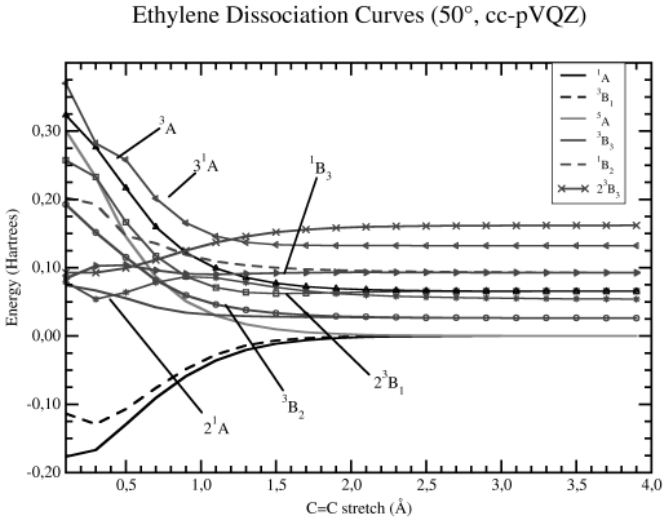


Figure 47: Dissociation curves for ethylene at $50°$ twist configuration computed with cc-pVQZ basis set. The graph shows the twelve selected states, those with lines in blue are the ones that have non-zero spin orbit coupling with the three lowest lying states at dissociation. The only state that doesn't couples is the triplet ^3A state.

C_2H_4 dissociation SOCC profiles.

The C_2H_4 dissociation SOCC profiles computed for the lowest energy states considered in principal text. The participation of the additional states at $50°$ twist geometry has the consequence that the SOCCs for the individual couplings acquire lower values than those obtained in the $0°$ case, i.e. the overall SOCC is distributed in more states.

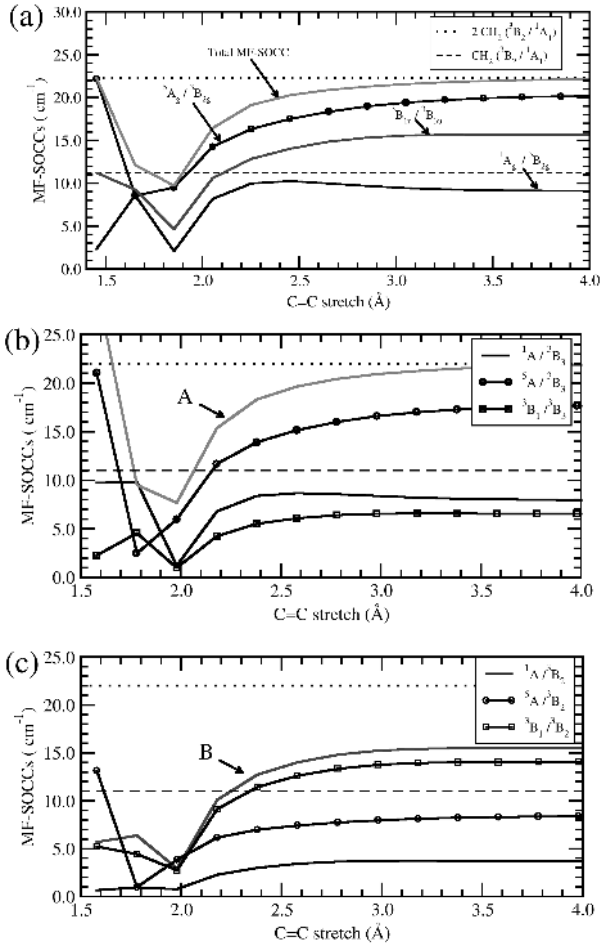


Figure 48: SOCC profiles for ethylene at (a) $0°$ and (b, c) $50°$ twist configurations along the C=C separation (RAS-SF/cc-pVQZ). The value of the MF-SOCC for the $^3B_2/^1A_1$ coupling from the CH_2 fragment is represented by the black-dashed line, the double of this quantity is indicated by the black-dotted line. Total SOCC is represented by the red curve.

Leading configurations for states of interest. In Table 30 we show the leading configurations in the RASCI wave function for the states involved in a non-zero SOC at dissociation.

Table 30: Selected states and dominant electronic configurations along dissociation at $0°$ and $50°$ torsion cases.

State	Leading Configuration
$0°$ twist	
^1A_g	$a_g^2 b_{1u}^2 b_{2g}^0 b_{3u}^0$
$^3B_{3u}$	$a_g^2 b_{1u}^1 b_{2g}^1 b_{3u}^0$
^5A_g	$a_g^1 b_{1u}^1 b_{2g}^1 b_{3u}^1$
$^3B_{2g}$	$a_g^2 b_{1u}^1 b_{2g}^0 b_{3u}^1$
$^3B_{1u}$	$a_g^2 b_{1u}^0 b_{2g}^1 b_{3u}^1$
$50°$ twist	
^1A	$a^2 b_2^2 b_3^0 b_1^0$
$^3B_1$	$a^2 b_2^1 b_3^1 b_1^0$
^5A	$a^1 b_2^1 b_3^1 b_1^1$
$^3B_3$	$a^2 b_2^1 b_3^0 b_1^1$
$^3B_2$	$a^2 b_2^0 b_3^1 b_1^1$

Squared amplitudes for the triplet $^3B_{1u}$

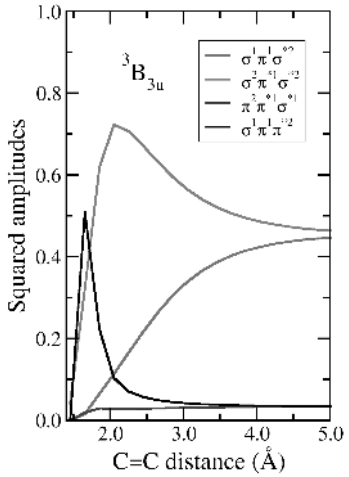


Figure 49: Plot of squared amplitudes values for ethylene's $^3B_{3u}$ state along C=C separation (RAS-SF/cc-pVQZ).

Natural transition orbitals and participation ratios for C_2H_4 at 50°

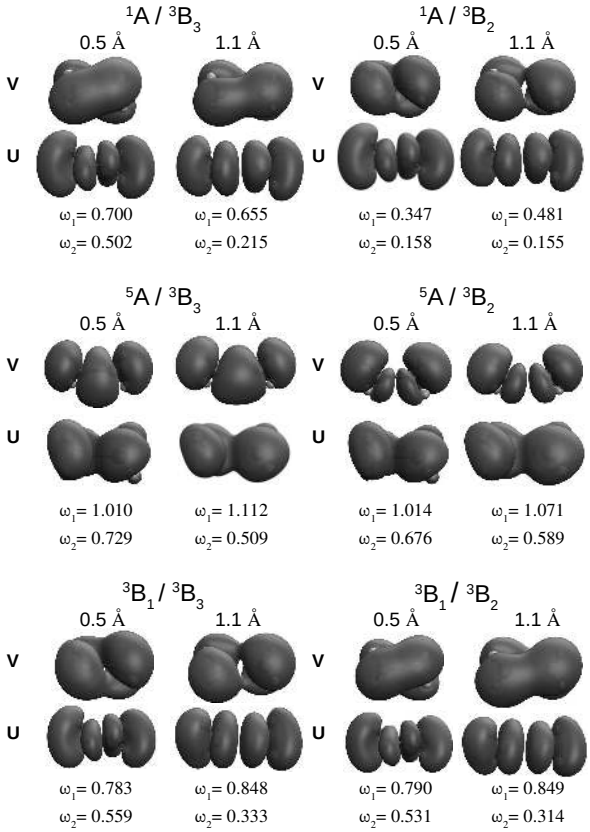


Figure 50: First dominant NTO pairs of the spinless transition matrix of ethylene (RAS-SF/cc-pVQZ) transitions at different stretch distances (isovalues of 0.05). The two largest singular values (ω_1 and ω_2) between the NTO pairs are presented below each pair.

Regarding the ${}^1A_g/{}^3B_{2g}$ and ${}^3B_{1u}/{}^3B_{3u}$ couplings, at ~ 1.85 Å (0.5 angs of stretch), the change in the character of the excitation is also accompanied with the suppression/enhancement of the SOCC. The PRNTOs at 1.65 Å (1.85 Å) is $\sim$2 ($\sim$1.7), both the $\pi^* \rightarrow \sigma$ and $\sigma^* \rightarrow \pi$ orbital excitations have an important weight in the transition. As the energy of the σ orbital is always lower than that for the π (see Figure 44) and the coupling is dominated by the $\pi^* \rightarrow \sigma$ transition, the SOCC increases due to the proximity of the σ orbital to nuclei (indicated indirectly by its lower orbital energy), the one-electron terms in SOC Hamiltonian are inversely proportional to $\sim \frac{Z_I}{r_{i,I}^3}$, and thus can be responsible of the SOC enhancement. At D region the transition is mostly dominated by the single excitation $\sigma^* \rightarrow \pi$, that at dissociation resembles the orbital transition involved in the ${}^3B_2/{}^1A_1$ coupling for the CH_2 fragments.

Complementary singular values and NTO-participation computed for C_2H_4 at 50 degrees dissociation.

Table 31: Singular values and participation rations for the leading NTO pairs for the non-analogous couplings that appear at 50° twist configuration.

C=C (Å)	ω_1	ω_2	PRNTOs
coupling: $^1A/^3B_1$			
0.3	0.3184	0.1527	1.6078
0.5	0.3469	0.1580	1.8520
1.1	0.4815	0.1556	1.4007
3.9	0.9395	0.3869	2.3283
coupling: $^5A/^3B_1$			
0.3	1.0510	0.6980	1.4339
0.5	1.0140	0.6763	2.0789
1.1	1.0710	0.5889	1.8772
3.9	0.8791	0.8720	2.3207
coupling: $^3B_1/^3B_3$			
0.3	0.8223	0.5202	1.7241
0.5	0.7826	0.5588	1.8709
1.1	0.8479	0.3330	1.5622
3.9	0.6782	0.6737	2.3288

Table 32: Singular values and participation rations for the leading NTO pairs couplings with higher energy states.

	0° planar			50° twist		
C=C (Å)	ω_1	ω_2	PRNTOs	ω_1	ω_2	PRNTOs
	coupling: $^1A_g/2^3B_{2g}$			coupling: $^1A/2^3B_3$		
0.3	0.7925	0.4858	1.6757	0.8024	0.3921	1.4747
0.5	0.7540	0.4848	1.7354	0.6924	0.4937	1.8464
1.1	0.6958	0.2099	1.3107	0.6489	0.2097	1.3663
3.9	0.3925	0.3872	2.2708	0.3914	0.3865	2.2939
	coupling: $^5A_g/2^3B_{2g}$			coupling: $^5A/2^3B_3$		
0.3	1.0040	0.6442	1.7309	1.0626	0.5311	1.5005
0.5	1.0450	0.6627	1.7525	1.0081	0.7152	1.8764
1.1	1.1301	0.4431	1.5355	1.1137	0.4687	1.6128
3.9	0.8752	0.8681	2.2710	0.8730	0.8665	2.2942
	coupling: $^3B_{3u}/^1B_{1u}$			coupling: $^3B_1/^1B_2$		
0.3	1.1508	0.3333	1.1788	1.1051	0.3989	1.2907
0.5	0.8328	0.6771	2.3795	0.8952	0.5608	2.3256
1.1	0.9866	0.3883	1.5139	0.8260	0.2587	1.2485
3.9	0.0060	0.0058	2.2336	0.0060	0.0059	2.2468

SOC in the ethylene-O_2 system

In the following data, we use the default Q-Chem symmetry labels. Here, the states ^3B, ^3A, ^1A, 2^1A and 3^1A are the corresponding X^3B, $^3(CT)$, $^1A(CT)$, $^1A(\Delta)$ and $^1A(\Sigma)$

discussed in the main text.

Table 33: Leading configurations and Head Gordon's N_u values at 5° and 55° of O_2 rotation (2.00 Å of separation), for higher energy states in order of magnitude of the RASCI weight.

^1B	$7b^1 9a^2 10a^1 + 7b^1 9a^1 10a^2$	2.00	2.00
2^1B	$7b^1 9a^1 10a^2 + 7b^1 9a^2 10a^1$	2.00	2.00

Along the molecular distortion, the lowest-lying states correspond to the different electronic occupation of the frontier π-orbitals formed as combinations of the C_2H_4 HOMO (labeled as $9a$) and the two singly occupied MOs (SOMOs) in ground state of O_2 ($10a$ and $7b$). Calculations indicate that the ^3B state (C_2 symmetry) is the lowest energy level for large torsion angles, with ^{1}A and ^{3}A states at higher energies. This state ordering changes upon coplanar rotation, with an avoided crossing between the two excited ^{1}A states at $\sim 50^\circ$. Ordering changes geometric configurations and particularly, it's coupling with ^1A states displays a large SOCC.

Table 34: Vertical excitation energies (in eV), of the lowest energy states for the EtO1 system under rotation of the O_2 fragment (RAS-SF/cc-pVQZ calculations). Only illustrative data was chosen.

angle	low-lying states				
0°	$^3A_2$	$^1A_2$	$^3B_2$	2^3B_2	$^1B_2$
	0.000	0.111	0.122	0.729	1.251
5°	^3A	^3B	^1A	2^1A	^1B
	0.000	0.110	0.109	1.197	1.239
25°	^3B	^1A	^3A	2^1A	^1B
	0.000	0.261	0.196	1.113	1.113
45°	^3B	^1A	^3A	^1B	2^1A
	0.000	0.810	0.798	1.095	1.121
65°	^3B	^1A	^1B	^3A	2^1A
	0.000	1.499	1.125	1.481	1.115
85°	^3B	^1A	^1B	^3A	2^1A
	0.000	1.989	1.172	1.939	1.127
90°	$^3B_1$	$^1A_2$	$^1B_2$	$^3A_2$	2^1A_2
	0.000	1.127	1.173	1.962	2.018

We focus on the description of ^3A, ^3B, ^1A and 2^1A states, an particularly on the SOC between ^3B with ^1A states. These are labeled as their order of apparition at 5° of torsion, respective leading configurations to the wave function (based on the RASCI weights) are illustrated in Table 35. As energy profiles show, at the beginning of the rotation (5° – 15°), a ^3A state is the lowest energy state, however the situation changes at ∼20° degrees, where the energy of the ^3A state is increased while that for the ^3B state is decreased becoming the lowest-energy state.

Along rotation, the ^1A state energy closely follows that for the ^3A tending to increase with the rotation angle. Obtained energy profiles predict a crossing near 55° rotation, involving the ^1A and 2^1A states, at this point there occurs a suppression/enhancement effect in the SOCC for the respective couplings with the ^3B state. Due to its larger energy (> 7.6 eV), a higher energy 3^1A state is not considered here. According with participation rations, the

^3B transitions acquire a multiconfigurational character that mainly involves the $7b \rightarrow 9a$ leading orbital transition for the $^3B/^1A$ coupling and the $7b \rightarrow 10a$ for the $^3B/2^1A$ case.

Table 35: Leading configurations and Head Gordon's N_u values at 5° and 55° of O_2 rotation (2.00 Å of separation), for the target states in order of magnitude of the RASCI weight.

State	Leading Configurations	N_u (5°)	N_u (55°)
^3A	$7b^2 9a^1 10a^1$	2.03	2.05
$3B$	$7b^1 9a^2 10a^1$	2.00	2.00
^1A	$7b^2 9a^1 10a^1 + 7b^2 9a^2 10a^0$	2.02	2.32
2^1A	$7b^0 9a^2 10a^2 + 7b^2 9a^0 10a^2 + 7b^2 9a^2 10a^0$	2.02	1.98
3^1A	$7b^2 9a^2 10a^0 + 7b^2 9a^0 10a^2 + 7b^0 9a^2 10a^2$	2.09	1.67

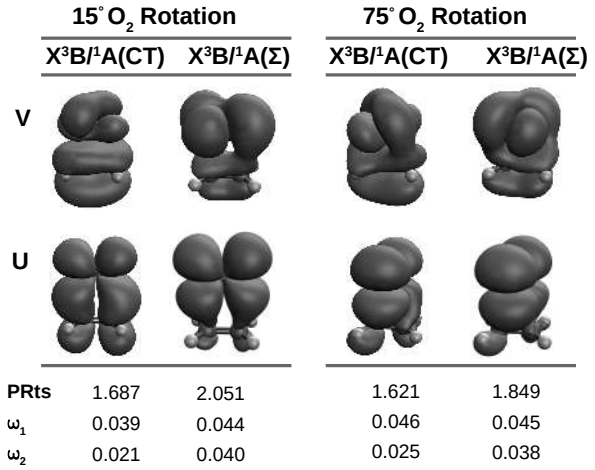


Figure 51: Dominant NTO pairs of the spinless transition matrix of $C_2H_4 \cdots O_2$ configuration involved in the $^3B \leftrightarrow^1 A$ and $^3B \leftrightarrow 3^1A$ transitions near the crossing regions, at 2.4 Å (15° rotation) and 2.1 Å (75° rotation) separations of O_2 fragment, with isovalues of 0.05. The participation ratios (PRts) and the two largest singular values (ω_1 and ω_2) between the NTO pairs are presented below each case.

List of Publications

José Aarón Rodríguez-Jiménez, Abel Carreras, David Casanova
J. Chem. Phys. 154, 124116 (2021); see chapter 3
Short-range DFT energy correction to multiconfigurational wave functions for open-shell systems

José Aarón Rodríguez-Jiménez, Abel Carreras, David Casanova
J. Phys. Chem. A 127, 5, 1206−1218 (2023); see chapter 6
Spin−Orbit Couplings of Open-Shell Systems with Restricted Active Space Configuration Interaction

José Aarón Rodríguez-Jiménez, Abel Carreras, David Casanova
J. Chem. Theory Comput. 20, 3, 1118-1129 (2024); see chapter 4
Small-occupation density functional correlation energy correction to wave function approximations.

Bibliography

[1] John A Pople. Nobel lecture: Quantum chemical models. *Reviews of Modern Physics*, 71(5):1267, 1999. See pages xv, 2, 9, and 10.

[2] Klaus-Peter Huber. *Molecular spectra and molecular structure: IV. Constants of diatomic molecules*. Springer Science & Business Media, 2013. See pages xxi, 117, 217, 218, and 219.

[3] Subhas J Chakravorty, Steven R Gwaltney, Ernest R Davidson, Farid A Parpia, and Charlotte Froese p Fischer. Ground-state correlation energies for atomic ions with 3 to 18 electrons. *Physical Review A*, 47(5):3649, 1993. See pages xxv, 166.

[4] K. P. Huber and G. Herzberg. *Constants of diatomic molecules*, pages 8–689. Springer US, Boston, MA, 1979. See pages xxv, 96, 98, and 170.

[5] Peter J Knowles and Nicholas C Handy. A new determinant-based full configuration interaction method. *Chemical physics letters*, 111(4-5):315–321, 1984. See pages 2, 28, and 83.

[6] Jeppe Olsen, Björn O Roos, Poul Jo/rgensen, and Hans Jo/rgen Aa Jensen. Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *The Journal of chemical physics*, 89(4):2185–2192, 1988. See pages 2, 28, 83, and 89.

[7] Jeppe Olsen, Poul Jørgensen, and Jack Simons. Passing the one-billion limit in full configuration-interaction (fci) calculations. *Chemical Physics Letters*, 169(6):463–472, 1990. See pages 2, 83.

[8] Pierre Hohenberg and Walter Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964. See pages 2, 29, and 85.

[9] Walter Kohn and Lu Jeu Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133, 1965. See pages 2, 30, 31, and 60.

[10] Marcel Nooijen and Rodney J Bartlett. Equation of motion coupled cluster method for electron attachment. *The Journal of chemical physics*, 102(9):3629–3647, 1995. See page 2.

[11] Jimmie D Doll and William P Reinhardt. Many-body green's functions for finite, nonuniform systems: Applications to closed shell atoms. *The Journal of Chemical Physics*, 57(3):1169–1184, 1972. See page 2.

[12] JV Ortiz. Dyson-orbital concepts for description of electrons in molecules. *The Journal of chemical physics*, 153(7), 2020. See page 2.

[13] Erica Mitchell, Thais Scott, Jie J Bao, Laura Gagliardi, and Donald Truhlar. Multiconfiguration pair-density functional theory and its application to the diels-alder reaction. *American Chemical Society SciMeetings*, 1(1), 2020. See page 2.

[14] Jack Simons. Why is quantum chemistry so complicated? *Journal of the American Chemical Society*, 145(8):4343–4354, 2023. See pages 2, 19.

[15] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004. See pages 2, 3, 70, and 84.

[16] Attila Szabo and Neil S Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012. See pages 2, 39.

[17] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2013. See pages 2, 5, 6, 11, 27, and 54.

[18] Philip Jakobsen and Frank Jensen. Representing exact electron densities by a single slater determinant in finite basis sets. *Journal of Chemical Theory and Computation*, 17(1):269–276, 2020. See page 3.

[19] David P Tew, Wim Klopper, and Trygve Helgaker. Electron correlation: The many-body problem at the heart of chemistry. *Journal of computational chemistry*, 28(8):1307–1320, 2007. See page 3.

[20] Christof Hattig, Wim Klopper, Andreas Kohn, and David P Tew. Explicitly correlated electrons in molecules. *Chemical reviews*, 112(1):4–74, 2012. See page 3.

[21] Klaus Capelle and Vivaldo L Campo Jr. Density functionals and model hamiltonians: Pillars of many-particle physics. *Physics Reports*, 528(3):91–159, 2013. See pages 4, 84.

[22] Andreas Savin. A combined density functional and configuration interaction method. *International Journal of Quantum Chemistry*, 34(S22):59–69, 1988. See pages 4, 8, 57, 84, 87, and 88.

[23] Andreas Savin. *Correlation Contributions from Density Functionals*, pages 213–230. Springer-Verlag, 1991. See pages 4, 57, 84, 87, 88, and 93.

[24] Björn O Roos, Mikołaj Szulkin, and Michał Jaszuński. Dynamic correlation for mcscf wave functions: An effective potential method. *Theoretica chimica acta*, 71:375–384, 1987. See pages 4, 57, 102, and 106.

[25] R McWeeny. Present status of the correlation problem. In *The New World of Quantum Chemistry: Proceedings of the Second International Congress of Quantum Chemistry Held at New Orleans, USA, April 19–24, 1976*, pages 3–31. Springer, 1976. See pages 4, 101, 102, 104, and 107.

[26] Egil A Hylleraas. The schrödinger two-electron atomic problem. In *Advances in Quantum Chemistry*, volume 1, pages 1–33. Elsevier, 1964. See pages 4, 57, and 102.

[27] Toshikatsu Koga. Hylleraas wave functions revisited. *The Journal of chemical physics*, 96(2):1276–1279, 1992. See pages 4, 57, and 102.

[28] Toshikatsu Koga and Kazunori Matsui. Optimal hylleraas wave functions. *Zeitschrift für Physik D Atoms, Molecules and Clusters*, 27:97–102, 1993. See pages 4, 57, and 102.

[29] Hubert M James and Albert Sprague Coolidge. The ground state of the hydrogen molecule. *The Journal of Chemical Physics*, 1(12):825–835, 1933. See pages 4, 102.

[30] Chaim L Pekeris. Ground state of two-electron atoms. *Physical Review*, 112(5):1649, 1958. See pages 4, 102.

[31] Wlodzimierz Kolos and Lutoslaw Wolniewicz. Improved theoretical ground-state energy of the hydrogen molecule. *The Journal of chemical physics*, 49(1):404–410, 1968. See pages 4, 102.

[32] W Heisenberg. Zur theorie des ferromagnetismus. *Z. Phys*, 49:619, 1928. See page 4.

[33] Paul Adrien Maurice Dirac. Quantum mechanics of many-electron systems. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 123(792):714–733, 1929. See page 4.

[34] W PAULI Jr. On the quantum mechanics of magnetic electrons. *Nature*, 119:282, 1927. See pages 4, 134.

[35] E Wigner and F Seitz. On the constitution of metallic sodium. ii. *Physical Review*, 46(6):509, 1934. See page 5.

[36] Per-Olov Löwdin. Quantum theory of many-particle systems. i. physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Physical Review*, 97(6):1474, 1955. See pages 5, 48, 49, 50, and 166.

[37] Krishnan Raghavachari, Gary W Trucks, John A Pople, and Martin Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chemical Physics Letters*, 157(6):479–483, 1989. See pages 5, 7, and 84.

[38] Ernest R Davidson, Stanley A Hagstrom, Subhas J Chakravorty, Verena Meiser Umar, and Charlotte Froese Fischer. Ground-state correlation energies for two-to ten-electron atomic ions. *Physical Review A*, 44(11):7071, 1991. See page 5.

[39] JE Lennard-Jones. The spatial correlation of electrons in molecules. *The Journal of Chemical Physics*, 20(6):1024–1029, 1952. See page 5.

[40] MA Buijse and EJ Baerends. Fermi holes and coulomb holes. *Density Functional Theory of Molecules, Clusters, and Solids*, pages 1–46, 1996. See pages 5, 54.

[41] Werner Kutzelnigg, Giuseppe Del Re, and Gaston Berthier. Correlation coefficients for electronic wave functions. *Physical Review*, 172(1):49, 1968. See pages 5, 6.

[42] Dieter Cremer. Density functional theory: coverage of dynamic and non-dynamic electron correlation effects. *Molecular Physics*, 99(23):1899–1940, 2001. See pages 5, 6.

[43] Oktay Sinanoğlu. Many-electron theory of atoms, molecules and their interactions. *Advances in Chemical Physics*, pages 315–412, 1964. See pages 6, 85.

[44] Aron J Cohen and Nicholas C Handy. Dynamic correlation. *Molecular Physics*, 99(7):607–615, 2001. See page 6.

[45] Daniel KW Mok, Ralf Neumann, and Nicholas C Handy. Dynamical and nondynamical correlation. *The Journal of Physical Chemistry*, 100(15):6225–6230, 1996. See pages 6, 50.

[46] Eloy Ramos-Cordoba, Pedro Salvador, and Eduard Matito. Separation of dynamic and nondynamic correlation. *Physical Chemistry Chemical Physics*, 18(34):24015–24023, 2016. See pages 6, 50, 51, and 85.

[47] E Valderrama, EV Ludeña, and J Hinze. Analysis of dynamical and nondynamical components of electron correlation energy by means of local-scaling density-functional theory. *The Journal of Chemical Physics*, 106(22):9227–9235, 1997. See pages 6, 93.

[48] E Valderrama, EV Ludena, and J Hinze. Assessment of dynamical and nondynamical correlation energy components for the beryllium-atom isoelectronic sequence. *The Journal of chemical physics*, 110(5):2343–2353, 1999. See page 6.

[49] Jerzy Cioslowski. Density-driven self-consistent-field method: Density-constrained correlation energies in the helium series. *Physical Review A*, 43(3):1223, 1991. See pages 6, 93.

[50] R. G. Parr and W. Yang. *Density-functional theory of atoms and molecules*. Oxford University Press: New York, 1989. See page 6.

[51] E. K. U. Gross, M. Petersilka, and T. Grabo. *Conventional Quantum Chemical Correlation Energy Versus Density-Functional Correlation Energy*. American Chemical Society, 1996. See page 6.

[52] Chr Møller and Milton S Plesset. Note on an approximation treatment for many-electron systems. *Physical review*, 46(7):618, 1934. See pages 6, 7.

[53] Isaiah Shavitt. The method of configuration interaction. In *Methods of electronic structure theory*, pages 189–275. Springer, 1977. See pages 6, 17, and 19.

[54] Jiří Čížek. On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods. *The Journal of Chemical Physics*, 45(11):4256–4266, 1966. See page 6.

[55] Robert J Buenker and Sigrid D Peyerimhoff. Individualized configuration selection in ci calculations with subsequent energy extrapolation. *Theor. Chim. Acta*, 35(1):33–58, 1974. See pages 6, 60.

[56] Robert J Buenker, Sigrid D Peyerimhoff, and Werner Butscher. Applicability of the multi-reference double-excitation ci (mrd-ci) method to the calculation of electronic wavefunctions and comparison with related techniques. *Mol. Phys.*, 35(3):771–791, 1978. See pages 6, 60.

[57] Hans-Joachim Werner and Peter J Knowles. An efficient internally contracted multiconfiguration–reference configuration interaction method. *The Journal of chemical physics*, 89(9):5803–5814, 1988. See page 6.

[58] Björn O Roos, Peter R Taylor, and Per EM Sigbahn. A complete active space scf method (casscf) using a density matrix formulated super-ci approach. *Chemical Physics*, 48(2):157–173, 1980. See pages 6, 12, 27, 60, 84, and 85.

[59] Björn O Roos, Per Linse, Per EM Siegbahn, and Margareta RA Blomberg. A simple method for the evaluation of the second-order-perturbation energy from external double-excitations with a casscf reference wavefunction. *Chemical Physics*, 66(1-2):197–207, 1982. See pages 6, 28, 60, and 84.

[60] Frank Neese, Frank Wennmohs, and Andreas Hansen. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *The Journal of chemical physics*, 130(11), 2009. See page 7.

[61] Christoph Riplinger, Barbara Sandhoefer, Andreas Hansen, and Frank Neese. Natural triple excitations in local coupled cluster calculations with pair natural orbitals. *The Journal of chemical physics*, 139(13), 2013. See pages 7, 12.

[62] Zoltán Rolik, Lóránt Szegedy, István Ladjánszki, Bence Ladóczki, and Mihály Kállay. An efficient linear-scaling ccsd (t) method based on local natural orbitals. *The Journal of chemical physics*, 139(9), 2013. See page 7.

[63] Max Schwilk, Qianli Ma, Christoph Köppl, and Hans-Joachim Werner. Scalable electron correlation methods. 3. efficient and accurate parallel local coupled cluster with pair natural orbitals (pno-lccsd). *Journal of Chemical Theory and Computation*, 13(8):3650–3675, 2017. See page 7.

[64] Steven R White and David A Huse. Numerical renormalization-group study of low-lying eigenstates of the antiferromagnetic s= 1 heisenberg chain. *Physical Review B*, 48(6):3844, 1993. See page 7.

[65] Rodney J Bartlett and Monika Musiał. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, 79(1):291, 2007. See page 7.

[66] Dmitry I Lyakh, Monika Musiał, Victor F Lotrich, and Rodney J Bartlett. Multireference nature of chemistry: The coupled-cluster view. *Chemical reviews*, 112(1):182–243, 2012. See pages 7, 28.

[67] George D Purvis III and Rodney J Bartlett. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *The Journal of Chemical Physics*, 76(4):1910–1918, 1982. See pages 7, 84.

[68] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Challenges for density functional theory. *Chemical reviews*, 112(1):289–320, 2012. See pages 7, 84.

[69] Kyle R Bryenton, Adebayo A Adeleke, Stephen G Dale, and Erin R Johnson. Delocalization error: The greatest outstanding challenge in density-functional theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 13(2):e1631, 2023. See pages 7, 8.

[70] Ewa Pastorczak, Michał Hapka, Libor Veis, and Katarzyna Pernal. Capturing the dynamic correlation for arbitrary spin-symmetry casscf reference with adiabatic connection approaches: Insights into the electronic structure of the tetramethyleneethane diradical. *The Journal of Physical Chemistry Letters*, 10(16):4668–4674, 2019. See pages 8, 44.

[71] Joshua W Hollett and Pierre-François Loos. Capturing static and dynamic correlation with δno-mp2 and δno-ccsd. *The Journal of Chemical Physics*, 152(1), 2020. See page 8.

[72] Mario Piris. Global natural orbital functional: Towards the complete description of the electron correlation. *Physical Review Letters*, 127(23):233001, 2021. See pages 8, 28.

[73] Soumen Ghosh, Pragya Verma, Christopher J Cramer, Laura Gagliardi, and Donald G Truhlar. Combining wave function methods with density functional theory for excited states. *Chemical reviews*, 118(15):7249–7292, 2018. See pages 8, 29, and 84.

[74] Jürgen Gräfenstein and Dieter Cremer*. Development of a cas-dft method covering non-dynamical and dynamical electron correlation in a balanced way. *Molecular Physics*, 103(2-3):279–308, 2005. See pages 8, 37.

[75] Jürgen Gräfenstein and Dieter Cremer. The combination of density functional theory with multi-configuration methods–cas-dft. *Chemical Physics Letters*, 316(5-6):569–577, 2000. See pages 8, 36.

[76] Paola Gori-Giorgi and Andreas Savin. Properties of short-range and long-range correlation energy density functionals from electron-electron coalescence. *Phys. Rev. A*, 73(3):032506, 2006. See pages 8, 60, and 67.

[77] Andreas Savin. Models and corrections: Range separation for electronic interaction—Lessons from density functional theory. *The Journal of Chemical Physics*, 153(16):160901, 10 2020. See pages 8, 84.

[78] Thierry Leininger, Hermann Stoll, Hans-Joachim Werner, Andreas Savin, et al. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.*, 275(3):151–160, 1997. See pages 8, 43, and 60.

[79] Emmanuel Fromager, Julien Toulouse, and Hans Jørgen Aa Jensen. On the universality of the long-/short-range separation in multiconfigurational density-functional theory. *J. Chem. Phys.*, 126(7):074111, 2007. See pages 8, 60.

[80] Erich Goll, Hans-Joachim Werner, Hermann Stoll, Thierry Leininger, Paola Gori-Giorgi, and Andreas Savin. A short-range gradient-corrected spin density functional in combination with long-range coupled-cluster methods: Application to alkali-metal rare-gas dimers. *Chem. Phys.*, 329(1-3):276–282, 2006. See pages 8, 62.

[81] John A Pople, J Stephen Binkley, and Rolf Seeger. Theoretical models incorporating electron correlation. *International Journal of Quantum Chemistry*, 10(S10):1–19, 1976. See page 9.

[82] Rodney J Bartlett. Many-body perturbation theory and coupled cluster theory for electron correlation in molecules. *Annual review of physical chemistry*, 32(1):359–401, 1981. See pages 9, 10, and 12.

[83] Krishnan Raghavachari. Electron correlation techniques in quantum chemistry: Recent advances. *Annual Review of Physical Chemistry*, 42(1):615–642, 1991. See pages 9, 28.

[84] Markus Reiher. Molecule-specific uncertainty quantification in quantum chemical studies. *Israel Journal of Chemistry*, 62(1-2):e202100101, 2022. See page 9.

[85] Henrik Koch, Hans Jo/rgen Aa Jensen, Poul Jo/rgensen, and Trygve Helgaker. Excitation energies from the coupled cluster singles and doubles linear response function (ccsdlr). applications to be, ch+, co, and h2o. *The Journal of chemical physics*, 93(5):3345–3350, 1990. See page 11.

[86] Klaus Müller-Dethlefs and Pavel Hobza. Noncovalent interactions: a challenge for experiment and theory. *Chemical Reviews*, 100(1):143–168, 2000. See page 11.

[87] Charles W Bauschlicher and Harry Partridge. On the invariance of the configuration interaction energy with respect to orbital rotations. *Theoretica chimica acta*, 85:255–259, 1993. See page 12.

[88] Francesco A Evangelista and Jürgen Gauss. Insights into the orbital invariance problem in state-specific multireference coupled cluster theory. *The Journal of chemical physics*, 133(4), 2010. See page 12.

[89] Liguo Kong. Orbital invariance issue in multireference methods. *International Journal of Quantum Chemistry*, 110(14):2603–2613, 2010. See page 12.

[90] Lin Shen, Binbin Xie, Ziwen Li, Lihong Liu, Ganglong Cui, and Wei-Hai Fang. Role of multistate intersections in photochemistry. *The Journal of Physical Chemistry Letters*, 11(20):8490–8501, 2020. See page 12.

[91] Luca Nanni. A discussion on the relativistic corrections of the electronic structures of multi-electron atoms. *International Journal of Quantum Chemistry*, 123(19):e27182, 2023. See page 12.

[92] Rodney J Barlett, Hideo Sekino, and George D Purvis III. Comparison of mbpt and coupled-cluster methods with full ci. importance of triplet excitation and infinite summations. *Chemical physics letters*, 98(1):66–71, 1983. See page 12.

[93] Hendrik Ḿuller and Werner Kutzelnigg. A ccsd (t)-r12 study of the ten-electron systems ne, f-, hf, h2o, nh3, nh4+ and ch4. *Molecular Physics*, 92(3):535–546, 1997. See page 12.

[94] Jan Rezac, Lucia Simova, and Pavel Hobza. Ccsd [t] describes noncovalent interactions better than the ccsd (t), ccsd (tq), and ccsdt methods. *Journal of chemical theory and computation*, 9(1):364–369, 2013. See page 12.

[95] Hendrik J Monkhorst. Chemical physics without the born-oppenheimer approximation: The molecular coupled-cluster method. *Physical Review A*, 36(4):1544, 1987. See page 12.

[96] Susi Lehtola. A review on non-relativistic, fully numerical electronic structure calculations on atoms and diatomic molecules. *International Journal of Quantum Chemistry*, 119(19):e25968, 2019. See page 14.

[97] Volker Blum, Ralf Gehrke, Felix Hanke, Paula Havu, Ville Havu, Xinguo Ren, Karsten Reuter, and Matthias Scheffler. Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications*, 180(11):2175–2196, 2009. See page 14.

[98] Igor Ying Zhang, Xinguo Ren, Patrick Rinke, Volker Blum, and Matthias Scheffler. Numeric atom-centered-orbital basis sets with valence-correlation consistency from h to ar. *New Journal of Physics*, 15(12):123033, 2013. See page 14.

[99] Kurt Lejaeghere, Gustav Bihlmayer, Torbjörn Björkman, Peter Blaha, Stefan Blügel, Volker Blum, Damien Caliste, Ivano E Castelli, Stewart J Clark, Andrea Dal Corso, et al. Reproducibility in density functional theory calculations of solids. *Science*, 351(6280):aad3000, 2016. See page 14.

[100] Stig Rune Jensen, Santanu Saha, José A Flores-Livas, William Huhn, Volker Blum, Stefan Goedecker, and Luca Frediani. The elephant in the room of density functional theory calculations. *The journal of physical chemistry letters*, 8(7):1449–1457, 2017. See page 14.

[101] Ryoichi Kikuchi. Gaussian functions in molecular integrals. *The Journal of Chemical Physics*, 22(1):148–148, 1954. See page 14.

[102] S Francis Boys. Electronic wave functions-i. a general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 200(1063):542–554, 1950. See pages 14, 107, 193, and 205.

[103] Samuel Francis Boys. The integral formulae for the variational solution of the molecular many-electron wave equation in terms of gaussian functions with direct electronic correlation. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 258(1294):402–411, 1960. See pages 14, 107, 193, and 207.

[104] Larry E McMurchie and Ernest R Davidson. One-and two-electron integrals over cartesian gaussian functions. *Journal of Computational Physics*, 26(2):218–231, 1978. See pages 14, 107.

[105] Shigeru Obara and A Saika. Efficient recursive computation of molecular integrals over cartesian gaussian functions. *The Journal of chemical physics*, 84(7):3963–3974, 1986. See page 14.

[106] Peter MW Gill and John A Pople. The prism algorithm for two-electron integrals. *International journal of quantum chemistry*, 40(6):753–772, 1991. See page 14.

[107] Mireia Güell, Josep M Luis, Miquel Sola, and Marcel Swart. Importance of the basis set for the spin-state energetics of iron complexes. *J. Phys. Chem. A*, 112(28):6384–6391, 2008. See page 14.

[108] Robertus van Leeuwen. *Kohn-Sham potentials in density functional theory*. Vrije Universiteit Amsterdam, The Netherlands, 1994. See pages 17, 29.

[109] István Mayer. *Simple theorems, proofs, and derivations in quantum chemistry*. Springer Science & Business Media, 2003. See page 18.

[110] L Meissner. Size-consistency corrections for configuration interaction calculations. *Chemical physics letters*, 146(3-4):204–210, 1988. See page 19.

[111] JML Martin, JP François, and R Gijbels. On size-consistency corrections for limited configuration-interaction calculations. *Chemical physics letters*, 172(5):346–353, 1990. See page 19.

[112] Bernard Huron, JP Malrieu, and P Rancurel. Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions. *The Journal of Chemical Physics*, 58(12):5745–5759, 1973. See page 20.

[113] Yann Garniron, Thomas Applencourt, Kevin Gasperich, Anouar Benali, Anthony Ferté, Julien Paquier, Barthélémy Pradines, Roland Assaraf, Peter Reinhardt, Julien Toulouse, et al. Quantum package 2.0: An open-source determinant-driven suite of programs. *Journal of chemical theory and computation*, 15(6):3591–3609, 2019. See pages 20, 92, and 117.

[114] Monika Dash, Saverio Moroni, Anthony Scemama, and Claudia Filippi. Perturbatively selected configuration-interaction wave functions for efficient geometry optimization in quantum monte carlo. *Journal of chemical theory and computation*, 14(8):4176–4182, 2018. See page 20.

[115] David Casanova. Restricted active space configuration interaction methods for strong correlation: Recent developments. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, page e1561, 2021. See pages 21, 52, 89, 134, and 136.

[116] Per Åke Malmqvist, Alistair Rendell, and Björn O Roos. The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *Journal of Physical Chemistry*, 94(14):5477–5482, 1990. See page 21.

[117] Dongxia Ma, Giovanni Li Manni, and Laura Gagliardi. The generalized active space concept in multiconfigurational self-consistent field methods. *The Journal of chemical physics*, 135(4), 2011. See page 21.

[118] Matthew R Hermes and Laura Gagliardi. Multiconfigurational self-consistent field theory with density matrix embedding: The localized active space self-consistent field method. *Journal of chemical theory and computation*, 15(2):972–986, 2019. See page 21.

[119] Francesco Aquilante, Jochen Autschbach, Rebecca K Carlson, Liviu F Chibotaru, Mickaël G Delcey, Luca De Vico, Ignacio Fdez. Galván, Nicolas Ferré, Luis Manuel Frutos, Laura Gagliardi, et al. Molcas 8: New capabilities for multiconfigurational quantum chemical calculations across the periodic table, 2016. See page 21.

[120] Konstantinos D Vogiatzis, Dongxia Ma, Jeppe Olsen, Laura Gagliardi, and Wibe A De Jong. Pushing configuration-interaction to the limit: Towards massively parallel mcscf calculations. *The Journal of chemical physics*, 147(18), 2017. See page 21.

[121] Daniel S Levine, Diptarka Hait, Norm M Tubman, Susi Lehtola, K Birgitta Whaley, and Martin Head-Gordon. Casscf with extremely large active spaces using the adaptive sampling configuration interaction method. *Journal of chemical theory and computation*, 16(4):2340–2354, 2020. See pages 21, 28.

[122] Benjamin G Levine, Andrew S Durden, Michael P Esch, Fangchun Liang, and Yinan Shu. Cas without scf—why to use casci and where to get the orbitals. *The Journal of Chemical Physics*, 154(9), 2021. See pages 22, 28.

[123] Rev McWeeny. Some recent advances in density matrix theory. *Reviews of Modern Physics*, 32(2):335, 1960. See page 23.

[124] Eugene S Kryachko and Eduardo V Ludeña. *Energy density functional theory of many-electron systems*, volume 4. Springer Science & Business Media, 2012. See pages 23, 49.

[125] David Casanova. Avoided crossings, conical intersections, and low-lying excited states with a single reference method: The restricted active space spin-flip configuration interaction approach. *The Journal of Chemical Physics*, 137(8):084105, 2012. See pages 26, 61, 90, 92, and 163.

[126] David Casanova. Efficient implementation of restricted active space configuration interaction with the hole and particle approximation. *Journal of computational chemistry*, 34(9):720–730, 2013. See pages 26, 61, and 92.

[127] David Casanova and Anna I Krylov. Spin-flip methods in quantum chemistry. *Physical Chemistry Chemical Physics*, 22(8):4326–4342, 2020. See page 26.

[128] Jeppe Olsen. The casscf method: A perspective and commentary. *International Journal of Quantum Chemistry*, 111(13):3267–3272, 2011. See page 27.

[129] Michael W Schmidt and Mark S Gordon. The construction and interpretation of mcscf wavefunctions. *Annual review of physical chemistry*, 49(1):233–266, 1998. See pages 27, 50.

[130] Zsuzsanna Tóth and Peter Pulay. Comparison of methods for active orbital selection in multiconfigurational calculations. *Journal of chemical theory and computation*, 16(12):7328–7341, 2020. See pages 27, 50.

[131] Laura Gagliardi, Donald G Truhlar, Giovanni Li Manni, Rebecca K Carlson, Chad E Hoyer, and Junwei Lucas Bao. Multiconfiguration pair-density functional theory: A new way to treat strongly correlated systems. *Accounts of chemical research*, 50(1):66–73, 2017. See pages 27, 38, 39, and 40.

[132] Peter J Knowles and Nicholas C Handy. A determinant based full configuration interaction program. *Computer physics communications*, 54(1):75–83, 1989. See page 28.

[133] S Zarrabian, CR Sarma, and J Paldus. Vectorizable approach to molecular ci problems using determinantal basis. *Chemical physics letters*, 155(2):183–188, 1989. See page 28.

[134] Gian Luigi Bendazzoli and Stefano Evangelisti. A vector and parallel full configuration interaction algorithm. *The Journal of chemical physics*, 98(4):3141–3150, 1993. See page 28.

[135] Peter J Knowles and Hans-Joachim Werner. An efficient second-order mc scf method for long configuration expansions. *Chemical physics letters*, 115(3):259–267, 1985. See page 28.

[136] Hans-Joachim Werner and Peter J Knowles. A second order multiconfiguration scf procedure with optimum convergence. *The Journal of chemical physics*, 82(11):5053–5063, 1985. See page 28.

[137] Todd J Martinez, Aseem Mehta, and Emily A Carter. Pseudospectral full configuration interaction. *The Journal of chemical physics*, 97(3):1876–1880, 1992. See page 28.

[138] Seiichiro Ten-no and Suehiro Iwata. Multiconfiguration self-consistent field procedure employing linear combination of atomic-electron distributions. *The Journal of chemical physics*, 105(9):3604–3611, 1996. See page 28.

[139] Francesco Aquilante, Thomas Bondo Pedersen, Roland Lindh, Björn Olof Roos, Alfredo Sánchez de Merás, and Henrik Koch. Accurate ab initio density fitting for multiconfigurational self-consistent field methods. *The Journal of chemical physics*, 129(2), 2008. See page 28.

[140] Lars Hedin. New method for calculating the one-particle green's function with application to the electron-gas problem. *Physical Review*, 139(3A):A796, 1965. See pages 28, 67.

[141] Lucia Reining. The gw approximation: content, successes and limitations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 8(3):e1344, 2018. See page 28.

[142] Mario Piris. Natural orbital functional theory. *Advances in Chemical Physics*, 134:387, 2007. See page 28.

[143] Mario Piris and Jesus M Ugalde. Perspective on natural orbital functional theory. *International Journal of Quantum Chemistry*, 114(18):1169–1175, 2014. See page 28.

[144] Pawel Tecmer and Katharina Boguslawski. Geminal-based electronic structure methods in quantum chemistry. toward a geminal model chemistry. *Physical Chemistry Chemical Physics*, 2022. See page 28.

[145] Mario Motta, Tanvi P Gujarati, Julia E Rice, Ashutosh Kumar, Conner Masteran, Joseph A Latone, Eunseok Lee, Edward F Valeev, and Tyler Y Takeshita. Quantum simulation of electronic structure with a transcorrelated hamiltonian: improved accuracy with a smaller footprint on the quantum computer. *Physical Chemistry Chemical Physics*, 22(42):24270–24281, 2020. See page 28.

[146] Ke Liao, Thomas Schraivogel, Hongjun Luo, Daniel Kats, and Ali Alavi. Towards efficient and accurate ab initio solutions to periodic systems via transcorrelation and coupled cluster theory. *Physical Review Research*, 3(3):033072, 2021. See page 28.

[147] Emmanuel Giner. A new form of transcorrelated hamiltonian inspired by range-separated dft. *The Journal of Chemical Physics*, 154(8), 2021. See page 28.

[148] Ke Liao, Huanchen Zhai, Evelin Martine Christlmaier, Thomas Schraivogel, Pablo López Ríos, Daniel Kats, and Ali Alavi. Density matrix renormalization group for transcorrelated hamiltonians: Ground and excited states in molecules. *Journal of Chemical Theory and Computation*, 19(6):1734–1743, 2023. See page 28.

[149] Kerstin Andersson, Per-Åke Malmqvist, and Björn O Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.*, 96(2):1218–1226, 1992. See pages 28, 60.

[150] Kerstin Andersson. Different forms of the zeroth-order hamiltonian in second-order perturbation theory with a complete active space self-consistent field reference function. *Theor. Chim. Acta*, 91(1-2):31–46, 1995. See pages 28, 60.

[151] David Casanova. Second-order perturbative corrections to the restricted active space configuration interaction with the hole and particle approach. *The Journal of Chemical Physics*, 140(14):144111, 2014. See pages 28, 61.

[152] Celestino Angeli, Renzo Cimiraglia, S Evangelisti, T Leininger, and J-P Malrieu. Introduction of n-electron valence states for multireference perturbation theory. *The Journal of Chemical Physics*, 114(23):10252–10264, 2001. See pages 28, 34, and 98.

[153] Celestino Angeli, Renzo Cimiraglia, and Jean-Paul Malrieu. n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *The Journal of chemical physics*, 117(20):9138–9153, 2002. See pages 28, 34, 60, and 98.

[154] Celestino Angeli, Mariachiara Pastore, and Renzo Cimiraglia. New perspectives in multireference perturbation theory: the n-electron valence state approach. *Theoretical Chemistry Accounts*, 117:743–754, 2007. See pages 28, 34.

[155] Josep M Bofill and Peter Pulay. The unrestricted natural orbital–complete active space (uno–cas) method: An inexpensive alternative to the complete active space–self-consistent-field (cas–scf) method. *The Journal of Chemical Physics*, 90(7):3637–3646, 1989. See pages 28, 50, and 84.

[156] Peter Pulay and Tracy P Hamilton. Uhf natural orbitals for defining and starting mc-scf calculations. *The Journal of Chemical Physics*, 88(8):4926–4933, 1988. See pages 28, 50.

[157] Sudip Chattopadhyay, Rajat K Chaudhuri, and Uttam Sinha Mahapatra. State-specific multireference perturbation theory with improved virtual orbitals: Taming the ground state of f 2, b e2, and n 2. *Journal of Computational Chemistry*, 36(12):907–925, 2015. See page 28.

[158] Suvonil Sinha Ray, Pradipta Ghosh, Rajat K Chaudhuri, and Sudip Chattopadhyay. Improved virtual orbitals in state specific multireference perturbation theory for prototypes of quasidegenerate electronic structure. *The Journal of Chemical Physics*, 146(6), 2017. See page 28.

[159] Kieron Burke. Perspective on density functional theory. *The Journal of Chemical Physics*, 136(15):150901, 04 2012. See pages 31, 86.

[160] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. Challenges for density functional theory. *Chem. Rev.*, 112(1):289–320, 2012. See pages 31, 60.

[161] Peter Pulay. A perspective on the caspt2 method. *International Journal of Quantum Chemistry*, 111(13):3273–3279, 2011. See page 32.

[162] Stefano Battaglia, Ignacio Fdez Galván, and Roland Lindh. Multiconfigurational quantum chemistry: The caspt2 method. In *Theoretical and Computational Photochemistry*, pages 135–162. Elsevier, 2023. See page 33.

[163] Roland Lindh and Ignacio Fdez Galván. Multi-configurational reference perturbation theory with a casscf reference function. *Quantum Chemistry and Dynamics of Excited States: Methods and Applications*, pages 299–353, 2020. See page 34.

[164] Stefano Battaglia, Lina Fransén, Ignacio Fdez. Galván, and Roland Lindh. Regularized caspt2: An intruder-state-free approach. *Journal of Chemical Theory and Computation*, 18(8):4814–4825, 2022. See page 34.

[165] Kenneth G Dyall. The choice of a zeroth-order hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function. *The Journal of chemical physics*, 102(12):4909–4918, 1995. See page 34.

[166] Niclas Forsberg and Per-Åke Malmqvist. Multiconfiguration perturbation theory with imaginary level shift. *Chemical Physics Letters*, 274(1-3):196–204, 1997. See page 34.

[167] Björn O Roos and Kerstin Andersson. Multiconfigurational perturbation theory with level shift—the cr2 potential revisited. *Chemical physics letters*, 245(2-3):215–223, 1995. See page 34.

[168] Celestino Angeli, Renzo Cimiraglia, and Jean-Paul Malrieu. N-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chemical physics letters*, 350(3-4):297–305, 2001. See pages 34, 35, and 98.

[169] Yang Guo, Kantharuban Sivalingam, and Frank Neese. Approximations of density matrices in n-electron valence state second-order perturbation theory (nevpt2). i. revisiting the nevpt2 construction. *The Journal of Chemical Physics*, 154(21), 2021. See page 35.

[170] Rudraditya Sarkar, Pierre-Francois Loos, Martial Boggio-Pasqua, and Denis Jacquemin. Assessing the performances of caspt2 and nevpt2 for vertical excitation energies. *Journal of Chemical Theory and Computation*, 18(4):2418–2436, 2022. See page 35.

[171] John P Perdew, Andreas Savin, and Kieron Burke. Escaping the symmetry dilemma through a pair-density interpretation of spin-density functional theory. *Physical Review A*, 51(6):4531, 1995. See pages 35, 38, 39, and 44.

[172] Burkhard Miehlich, Hermann Stoll, and Andreas Savin. A correlation-energy density functional for multideterminantal wavefunctions. *Molecular Physics*, 91(3):527–536, 1997. See pages 35, 36, 43, and 84.

[173] R Pollet, A Savin, T Leininger, and H Stoll. Combining multideterminantal wave functions with density functionals to handle near-degeneracy in atoms and molecules. *J. Chem. Phys.*, 116(4):1250–1258, 2002. See pages 37, 43, 44, and 60.

[174] Hermann Stoll. On the coupling of multi-configuration self-consistent-field and density-functional information. *Chemical physics letters*, 376(1-2):141–147, 2003. See page 37.

[175] R Takeda, S Yamanaka, and K Yamaguchi. Cas-dft based on odd-electron density and radical density. *Chemical physics letters*, 366(3-4):321–328, 2002. See page 37.

[176] K Nakata, T Ukai, S Yamanaka, T Takada, and K Yamaguchi. Casscf version of density functional theory. *International journal of quantum chemistry*, 106(15):3325–3333, 2006. See page 37.

[177] Shusuke Yamanaka, Kazuto Nakata, Takeshi Ukai, Toshikazu Takada, and Kizashi Yamaguchi. Multireference density functional theory with orbital-dependent correlation corrections. *International journal of quantum chemistry*, 106(15):3312–3324, 2006. See page 37.

[178] Christel M Marian, Adrian Heil, and Martin Kleinschmidt. The dft/mrci method. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 9(2):e1394, 2019. See page 37.

[179] Stefan Grimme and Mirko Waletzke. A combination of kohn–sham density functional theory and multi-reference configuration interaction methods. *The Journal of chemical physics*, 111(13):5645–5655, 1999. See pages 37, 38.

[180] Adrian Heil, Martin Kleinschmidt, and Christel M Marian. On the performance of dft/mrci hamiltonians for electronic excitations in transition metal complexes: The role of the damping function. *The Journal of Chemical Physics*, 149(16), 2018. See page 38.

[181] Igor Lyskov, Martin Kleinschmidt, and Christel M Marian. Redesign of the dft/mrci hamiltonian. *The Journal of chemical physics*, 144(3), 2016. See page 38.

[182] Adrian Heil and Christel M Marian. Dft/mrci hamiltonian for odd and even numbers of electrons. *The Journal of Chemical Physics*, 147(19), 2017. See page 38.

[183] Giovanni Li Manni, Rebecca K Carlson, Sijie Luo, Dongxia Ma, Jeppe Olsen, Donald G Truhlar, and Laura Gagliardi. Multiconfiguration pair-density functional theory. *Journal of chemical theory and computation*, 10(9):3669–3680, 2014. See pages 38, 39, 40, and 84.

[184] Prachi Sharma, Donald G Truhlar, and Laura Gagliardi. Active space dependence in multiconfiguration pair-density functional theory. *Journal of chemical theory and computation*, 14(2):660–669, 2018. See page 38.

[185] Olle Gunnarsson and Bengt I Lundqvist. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Physical Review B*, 13(10):4274, 1976. See pages 39, 44.

[186] Sergey Gusarov, Per-Åke Malmqvist, Roland Lindh, and Björn O Roos. Correlation potentials for a multiconfigurational-based density functional theory with exact exchange. *Theoretical Chemistry Accounts*, 112:84–94, 2004. See page 39.

[187] Chen Zhou, Matthew R Hermes, Dihua Wu, Jie J Bao, Riddhish Pandharkar, Daniel S King, Dayou Zhang, Thais R Scott, Aleksandr O Lykhin, Laura Gagliardi, et al. Electronic structure of strongly correlated systems: recent developments in multiconfiguration pair-density functional theory and multiconfiguration nonclassical-energy functional theory. *Chemical Science*, 13(26):7685–7706, 2022. See pages 40, 84.

[188] Oleg V Gritsenko, Robert Van Meer, and Katarzyna Pernal. Efficient evaluation of electron correlation along the bond-dissociation coordinate in the ground and excited ionic states with dynamic correlation suppression and enhancement functions of the on-top pair density. *Physical Review A*, 98(6):062510, 2018. See page 40.

[189] Michał Hapka, Katarzyna Pernal, and Oleg V Gritsenko. Local enhancement of dynamic correlation in excited states: Fresh perspective on ionicity and development of correlation density functional approximation based on the on-top pair density. *The journal of physical chemistry letters*, 11(15):5883–5889, 2020. See pages 40, 41.

[190] Oleg V Gritsenko and Katarzyna Pernal. Complete active space and corrected density functional theories helping each other to describe vertical electronic $\pi \rightarrow \pi^*$ excitations in prototype multiple-bonded molecules. *The Journal of Chemical Physics*, 151(2), 2019. See page 41.

[191] Michał Hapka, Katarzyna Pernal, and Oleg V Gritsenko. Molecular multibond dissociation with small complete active space augmented by correlation density functionals. *The Journal of Chemical Physics*, 152(20), 2020. See page 41.

[192] Andreas Savin and Heinz-Jürgen Flad. Density functionals for the Yukawa electron-electron interaction. *Int. J. Quantum Chem.*, 56(4):327–332, 1995. See pages 42, 61.

[193] Katarzyna Pernal and Michał Hapka. Range-separated multiconfigurational density functional theory methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 12(2):e1566, 2022. See page 42.

[194] Mel Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences*, 76(12):6062–6065, 1979. See pages 42, 85.

[195] Andreas Savin. Beyond the kohn–sham determinant. In *Recent Advances In Density Functional Methods: (Part I)*, pages 129–153. World Scientific, 1995. See page 43.

[196] Tosio Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Communications on Pure and Applied Mathematics*, 10(2):151–177, 1957. See pages 43, 104.

[197] Julien Toulouse. Review of approximations for the exchange-correlation energy in density-functional theory. In *Density Functional Theory: Modeling, Mathematical Analysis, Computational Methods, and Applications*, pages 1–90. Springer, 2022. See page 43.

[198] A. Savin. On degeneracy, near-degeneracy and density functional theory. In J.M. Seminario, editor, *Recent Developments and Applications of Modern Density Functional Theory*, volume 4 of *Theoretical and Computational Chemistry*, pages 327–357. Elsevier, 1996. See page 43.

[199] Hermann Stoll and Andreas Savin. Density functionals for correlation energies of atoms and molecules. In *Density functional methods in physics*, pages 177–207. Springer, 1985. See page 43.

[200] Andreas Savin and Heinz-Jürgen Flad. Density functionals for the yukawa electron-electron interaction. *Int. J. Quantum Chem.*, 56(4):327–332, 1995. See pages 43, 44, and 60.

[201] Kamal Sharkas, Andreas Savin, Hans Jørgen Aa Jensen, and Julien Toulouse. A multiconfigurational hybrid density-functional theory. *The Journal of Chemical Physics*, 137(4), 2012. See pages 43, 44.

[202] Emmanuel Fromager, Julien Toulouse, and Hans Jørgen Aa Jensen. On the universality of the long-/short-range separation in multiconfigurational density-functional theory. *J. Chem. Phys.*, 126(7):074111, 2007. See pages 44, 67.

[203] Emmanuel Fromager, Florent Réal, Pernilla Wåhlin, Ulf Wahlgren, and Hans Jørgen Aa Jensen. On the universality of the long-/short-range separation in multiconfigurational density-functional theory. ii. investigating f^0 actinide species. *J. Chem. Phys.*, 131(5):054107, 2009. See pages 44, 60.

[204] Erik Donovan Hedegård, Julien Toulouse, and Hans Jørgen Aagaard Jensen. Multiconfigurational short-range density-functional theory for open-shell systems. *The Journal of Chemical Physics*, 148(21), 2018. See page 44.

[205] Erik Donovan Hedegård, Frank Heiden, Stefan Knecht, Emmanuel Fromager, and Hans Jørgen Aagaard Jensen. Assessment of charge-transfer excitations with time-dependent, range-separated density functional theory based on long-range mp2 and multiconfigurational self-consistent field wave functions. *The Journal of chemical physics*, 139(18), 2013. See page 44.

[206] Emmanuel Fromager, Stefan Knecht, and Hans Jørgen Aa Jensen. Multi-configuration time-dependent density-functional theory based on range separation. *The Journal of Chemical Physics*, 138(8), 2013. See page 44.

[207] Erik Donovan Hedegård, Stefan Knecht, Jesper Skau Kielberg, Hans Jørgen Aagaard Jensen, and Markus Reiher. Density matrix renormalization group with efficient dynamical electron correlation through range separation. *The Journal of chemical physics*, 142(22):224108, 2015. See page 44.

[208] A Savin, F Colonna, and R Pollet. Adiabatic connection approach to density functional theory of electronic systems. *Int. J. Quantum Chem.*, 93(3):166–190, 2003. See pages 44, 60.

[209] David C Langreth and John P Perdew. The exchange-correlation energy of a metallic surface. *Solid State Communications*, 17(11):1425–1429, 1975. See page 44.

[210] Michael Seidl, Sara Giarrusso, Stefan Vuckovic, Eduardo Fabiano, and Paola Gori-Giorgi. Communication: Strong-interaction limit of an adiabatic connection in hartree-fock theory. *The Journal of chemical physics*, 149(24), 2018. See page 44.

[211] Timothy J Daas, Juri Grossi, Stefan Vuckovic, Ziad H Musslimani, Derk P Kooi, Michael Seidl, Klaas JH Giesbertz, and Paola Gori-Giorgi. Large coupling-strength expansion of the møller–plesset adiabatic connection: From paradigmatic cases to variational expressions for the leading terms. *The Journal of chemical physics*, 153(21), 2020. See page 44.

[212] Pavel Beran, Mikuláš Matoušek, Michał Hapka, Katarzyna Pernal, and Libor Veis. Density matrix renormalization group with dynamical correlation via adiabatic connection. *Journal of Chemical Theory and Computation*, 17(12):7575–7585, 2021. See page 44.

[213] Daria Drwal, Pavel Beran, Michał Hapka, Marcin Modrzejewski, Adam Sokół, Libor Veis, and Katarzyna Pernal. Efficient adiabatic connection approach for strongly correlated systems: Application to singlet–triplet gaps of biradicals. *The Journal of Physical Chemistry Letters*, 13(20):4570–4578, 2022. See page 44.

[214] Ewa Pastorczak and Katarzyna Pernal. Electronic excited states from the adiabatic-connection formalism with complete active space wave functions. *The Journal of Physical Chemistry Letters*, 9(18):5534–5538, 2018. See pages 44, 46.

[215] Daria Drwal, Ewa Pastorczak, and Katarzyna Pernal. Excited states in the adiabatic connection fluctuation-dissipation theory: Recovering missing correlation energy from the negative part of the density response spectrum. *The Journal of Chemical Physics*, 154(16), 2021. See page 44.

[216] Katarzyna Pernal. Electron correlation from the adiabatic connection for multireference wave functions. *Physical review letters*, 120(1):013001, 2018. See pages 44, 47.

[217] János Ángyán, John Dobson, Georg Jansen, and Tim Gould. *London dispersion forces in molecules, solids and nano-structures: an introduction to physical models and computational methods*. Royal Society of Chemistry, 2020. See page 45.

[218] Mikuláš Matoušek, Michał Hapka, Libor Veis, and Katarzyna Pernal. Toward more accurate adiabatic connection approach for multireference wavefunctions. *The Journal of Chemical Physics*, 158(5), 2023. See page 45.

[219] Koushik Chatterjee and Katarzyna Pernal. Excitation energies from extended random phase approximation employed with approximate one-and two-electron reduced density matrices. *The Journal of Chemical Physics*, 137(20), 2012. See page 46.

[220] Katarzyna Pernal, Koushik Chatterjee, and Piotr H Kowalski. How accurate is the strongly orthogonal geminal theory in predicting excitation energies? comparison of the extended random phase approximation and the linear response theory approaches. *The Journal of Chemical Physics*, 140(1), 2014. See page 46.

[221] Michał Hapka, Ewa Pastorczak, Agnieszka Krzemińska, and Katarzyna Pernal. Long-range-corrected multiconfiguration density functional with the on-top pair density. *The Journal of Chemical Physics*, 152(9), 2020. See page 47.

[222] Ernest R Davidson. Properties and uses of natural orbitals. *Reviews of Modern Physics*, 44(3):451, 1972. See pages 48, 49, 84, and 87.

[223] KJH Giesbertz and Robert van Leeuwen. Natural occupation numbers: When do they vanish? *The Journal of Chemical Physics*, 139(10):104109, 2013. See pages 48, 84.

[224] Ernest R Davidson. Configuration interaction description of electron correlation. In *The World of Quantum Chemistry: Proceedings of the First International Congress of Quantum Chemistry held at Menton, France, July 4–10, 1973*, pages 17–30. Springer, 1974. See page 49.

[225] F Illas, J Rubio, and JM Ricart. Approximate natural orbitals and the convergence of a second order multireference many-body perturbation theory (cipsi) algorithm. *The Journal of chemical physics*, 89(10):6376–6384, 1988. See page 49.

[226] KH Thunemann, J Römelt, SD Peyerimhoff, and RJ Buenker. A study of the convergence in iterative natural orbital procedures. *International Journal of Quantum Chemistry*, 11(5):743–752, 1977. See page 49.

[227] Andreas Grüneis, George H Booth, Martijn Marsman, James Spencer, Ali Alavi, and Georg Kresse. Natural orbitals for wave function based correlated calculations using a plane wave basis set. *Journal of chemical theory and computation*, 7(9):2780–2785, 2011. See page 49.

[228] James J Shepherd, Andreas Grüneis, George H Booth, Georg Kresse, and Ali Alavi. Convergence of many-body wave-function expansions using a plane-wave basis: From homogeneous electron gas to solid state systems. *Physical Review B*, 86(3):035111, 2012. See page 49.

[229] P Jeffrey Hay. On the calculation of natural orbitals by perturbation theory. *The Journal of Chemical Physics*, 59(5):2468–2476, 1973. See page 49.

[230] Per Jensen and PR Bunker. The potential surface and stretching frequencies of $\tilde{X}^3B_1$ methylene (CH_2) determined from experiment using the morse oscillator-rigid bender internal dynamics hamiltonian. *J. Chem. Phys.*, 89(3):1327–1332, 1988. See pages 50, 72, and 162.

[231] Charles Alfred Coulson and Inga Fischer. Xxxiv. notes on the molecular orbital treatment of the hydrogen molecule. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 40(303):386–393, 1949. See page 50.

[232] Christine Geron, George Dive, and Dominique Dehareng. The hartree–fock triplet instability: influence of conformation and substitution. *Journal of Molecular Structure: THEOCHEM*, 760(1-3):75–85, 2006. See page 50.

[233] Florian Krausbeck, David Mendive-Tapia, Alex JW Thom, and Michael J Bearpark. Choosing rasscf orbital active spaces for multiple electronic states. *Computational and Theoretical Chemistry*, 1040:14–19, 2014. See page 50.

[234] Mark S Gordon, Michael W Schmidt, Galina M Chaban, Kurt R Glaesemann, Walter J Stevens, and Carlos Gonzalez. A natural orbital diagnostic for multiconfigurational character in correlated wave functions. *The Journal of chemical physics*, 110(9):4199–4207, 1999. See page 50.

[235] Oksana Tishchenko, Jingjing Zheng, and Donald G Truhlar. Multireference model chemistries for thermochemical kinetics. *Journal of Chemical Theory and Computation*, 4(8):1208–1219, 2008. See page 50.

[236] Jerzy Cioslowski. Differential density matrix overlap: an index for assessment of electron correlation in atoms and molecules. *Theoretica chimica acta*, 81:319–327, 1992. See page 50.

[237] Eloy Ramos-Cordoba and Eduard Matito. Local descriptors of dynamic and nondynamic correlation. *Journal of Chemical Theory and Computation*, 13(6):2705–2711, 2017. See pages 50, 51.

[238] Xiang Xu, Luis Soriano-Agueda, Xabier López, Eloy Ramos-Cordoba, and Eduard Matito. All-purpose measure of electron correlation for multireference diagnostics. *Journal of Chemical Theory and Computation*, 2023. See page 50.

[239] Kazuo Takatsuka, Takayuki Fueno, and Kizashi Yamaguchi. Distribution of odd electrons in ground-state molecules. *Theoretica chimica acta*, 48:175–183, 1978. See pages 50, 52, and 136.

[240] Martin Head-Gordon. Characterizing unpaired electrons from the one-particle density matrix. *Chemical physics letters*, 372(3-4):508–511, 2003. See pages 50, 52, and 136.

[241] Viktor N Staroverov and Ernest R Davidson. Distribution of effectively unpaired electrons. *Chemical Physics Letters*, 330(1-2):161–168, 2000. See pages 50, 84.

[242] Curtis L Janssen and Ida MB Nielsen. New diagnostics for coupled-cluster and møller–plesset perturbation theory. *Chemical physics letters*, 290(4-6):423–430, 1998. See page 50.

[243] Róbert Izsák, Aleksei V Ivanov, Nick S Blunt, Nicole Holzmann, and Frank Neese. Measuring electron correlation: The impact of symmetry and orbital transformations. *Journal of Chemical Theory and Computation*, 2023. See page 50.

[244] Joshua W Hollett and Peter MW Gill. The two faces of static correlation. *The Journal of Chemical Physics*, 134(11), 2011. See pages 50, 93.

[245] Chenru Duan, Fang Liu, Aditya Nandy, and Heather J Kulik. Data-driven approaches can overcome the cost–accuracy trade-off in multireference diagnostics. *Journal of Chemical Theory and Computation*, 16(7):4373–4387, 2020. See page 50.

[246] Timothy J Lee and Peter R Taylor. A diagnostic for determining the quality of single-reference electron correlation methods. *International Journal of Quantum Chemistry*, 36(S23):199–207, 1989. See page 50.

[247] Uma R Fogueri, Sebastian Kozuch, Amir Karton, and Jan ML Martin. A simple dft-based diagnostic for nondynamical correlation. *Theoretical Chemistry Accounts*, 132:1–9, 2013. See page 50.

[248] Ida M Beck Nielsen and Curtis L Janssen. Double-substitution-based diagnostics for coupled-cluster and møller–plesset perturbation theory. *Chemical physics letters*, 310(5-6):568–576, 1999. See page 50.

[249] John S Sears and C David Sherrill. Assessing the performance of density functional theory for the electronic structure of metal- salens: The 3d0-metals. *The Journal of Physical Chemistry A*, 112(15):3466–3477, 2008. See page 50.

[250] Lu T Xu and Thom H Dunning. Dynamical electron correlation and the chemical bond. i. covalent bonds in ah and af (a= b–f). *The Journal of Chemical Physics*, 157(1), 2022. See page 50.

[251] Manoj K Kesharwani, Nitai Sylvetsky, Andreas Köhn, David P Tew, and Jan ML Martin. Do ccsd and approximate ccsd-f12 variants converge to the same basis set limits? the case of atomization energies. *The Journal of chemical physics*, 149(15), 2018. See page 51.

[252] Xing Xu. *Two Aspects of Electron Correlation: Multireference Diagnostics and London Dispersion Interactions*. Phd thesis, University of Basque Country UPV/EHU, Donostia-San Sebastian, Spain, November 2023. See pages 51, 108, and 125.

[253] Stefan Grimme and Andreas Hansen. A practicable real-space measure and visualization of static electron-correlation effects. *Angewandte Chemie International Edition*, 54(42):12308–12313, 2015. See page 51.

[254] Christoph Alexander Bauer, Andreas Hansen, and Stefan Grimme. The fractional occupation number weighted density as a versatile analysis tool for molecules with a complicated electronic structure. *Chemistry–A European Journal*, 23(25):6150–6164, 2017. See page 51.

[255] Natalie Orms, Dirk R Rehn, Andreas Dreuw, and Anna I Krylov. Characterizing bonding patterns in diradicals and triradicals by density-based wave function analysis: A uniform approach. *J. Chem. Theory Comput.*, 14(2):638–648, 2018. See pages 52, 74.

[256] Anna I Krylov. From orbitals to observables and back. *The Journal of Chemical Physics*, 153(8):080901, 2020. See pages 52, 53, 54, 136, and 149.

[257] Richard L Martin. Natural transition orbitals. *The Journal of chemical physics*, 118(11):4775–4777, 2003. See page 53.

[258] Stefanie A Mewes, Felix Plasser, Anna Krylov, and Andreas Dreuw. Benchmarking excited-state calculations using exciton properties. *Journal of Chemical Theory and Computation*, 14(2):710–725, 2018. See page 53.

[259] Elmer Valderrama, Jesus M Ugalde, and Russell J Boyd. Intracule and extracule densities: Historical perspectives and future prospects. *Many-Electron Densities and Reduced Density Matrices*, pages 231–248, 2000. See pages 54, 55, and 108.

[260] Ajit J Thakkar. Extracules, intracules, correlation holes, potentials, coefficients and all that. In *Density Matrices and Density Functionals: Proceedings of the A. John Coleman Symposium*, pages 553–581. Springer, 1987. See page 54.

[261] Roy McWeeny. *Methods of molecular quantum mechanics*. Academic-press London, 2nd edition, 1989. See page 54.

[262] E Valderrama, JM Mercero, and JM Ugalde. The separation of the dynamical and non-dynamical electron correlation effects. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 34(3):275, 2001. See page 55.

[263] Charles A Coulson. Present state of molecular structure calculations. *Reviews of Modern Physics*, 32(2):170, 1960. See pages 55, 108.

[264] CA Coulson and AH Neilson. Electron correlation in the ground state of helium. *Proceedings of the Physical Society*, 78(5):831, 1961. See pages 55, 108.

[265] RJ Boyd and CA Coulson. Coulomb hole in some excited states of helium. *Journal of Physics B: Atomic and Molecular Physics*, 6(5):782, 1973. See pages 55, 108.

[266] Mireia Via-Nadal, Mauricio Rodríguez-Mayorga, Eloy Ramos-Cordoba, and Eduard Matito. Natural range separation of the coulomb hole. *The Journal of chemical physics*, 156(18), 2022. See page 56.

[267] David Casanova. Short-range density functional correlation within the restricted active space ci method. *The Journal of Chemical Physics*, 148(12):124118, 2018. See pages 57, 60, 62, and 67.

[268] Abel Carreras, Hanjie Jiang, Pavel Pokhilko, Anna I Krylov, Paul M Zimmerman, and David Casanova. Calculation of spin–orbit couplings using rasci spinless one-particle density matrices: Theory and applications. *The Journal of Chemical Physics*, 153(21):214107, 2020. See pages 57, 58, 134, 136, and 139.

[269] Masayoshi Nakano and Benoît Champagne. Nonlinear optical properties in open-shell molecular systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 6(2):198–210, 2016. See pages 59, 60.

[270] Ganna Gryn'ova, Michelle L Coote, and Clemence Corminboeuf. Theory and practice of uncommon molecular electronic configurations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 5(6):440–459, 2015. See page 59.

[271] Markus Reiher and Alexander Wolf. *Relativistic quantum chemistry: the fundamental theory of molecular science*. John Wiley & Sons, 2014. See page 59.

[272] W. T. Borden, editor. *Diradicals*. Wiley, 1982. See page 59.

[273] Matthew S Platz and VM Maloney. *Kinetics and spectroscopy of carbenes and biradicals*. Springer, 1990. See page 59.

[274] George Christou, Dante Gatteschi, David N Hendrickson, and Roberta Sessoli. Single-molecule magnets. *MRS Bull*, 25(11):66–71, 2000. See page 60.

[275] Lionel Salem and Colin Rowland. The electronic properties of diradicals. *Angewandte Chemie International Edition*, 11(2):92–111, 1972. See pages 60, 71, 134, and 141.

[276] Thijs Stuyver, Bo Chen, Tao Zeng, Paul Geerlings, Frank De Proft, and Roald Hoffmann. Do diradicals behave like radicals? *Chem. Rev.*, 119(21):11291–11351, 2019. PMID: 31593450. See page 60.

[277] Xiaoguang Hu, Wenxiang Wang, Dongsheng Wang, and Yonghao Zheng. The electronic applications of stable diradicaloids: present and future. *J. Mater. Chem. C*, 6(42):11232–11242, 2018. See page 60.

[278] Thijs Stuyver, Tao Zeng, Yuta Tsuji, Paul Geerlings, and Frank De Proft. Diradical character as a guiding principle for the insightful design of molecular nanowires with an increasing conductance with length. *Nano Lett.*, 18(11):7298–7304, 2018. See page 60.

[279] Yuta Tsuji, Roald Hoffmann, Mikkel Strange, and Gemma C Solomon. Close relation between quantum interference in molecular conductance and diradical existence. *Proc. Natl. Acad. Sci. U.S.A.*, 113(4):E413–E419, 2016. See page 60.

[280] Hiroshi Nagai, Masayoshi Nakano, Kyohei Yoneda, Ryohei Kishi, Hideaki Takahashi, Akihiro Shimizu, Takashi Kubo, Kenji Kamada, Koji Ohta, Edith Botek, et al. Signature of multiradical character in second hyperpolarizabilities of rectangular graphene nanoflakes. *Chem. Phys. Let.*, 489(4-6):212–218, 2010. See page 60.

[281] Paul M Lahti. *Magnetic properties of organic materials*. CRC Press, 1999. See page 60.

[282] Millicent B Smith and Josef Michl. Singlet fission. *Chem. Rev.*, 110(11):6891–6936, 2010. See page 60.

[283] Millicent B Smith and Josef Michl. Recent advances in singlet fission. *Annu. Rev. Phys. Chem*, 64:361–386, 2013. See page 60.

[284] David Casanova. Theoretical modeling of singlet fission. *Chem. Rev.*, 118(15):7164–7207, 2018. See page 60.

[285] Vladimir V Zhivonitko, Jonas Bresien, Axel Schulz, and Igor V Koptyug. Parahydrogen-induced polarization with a metal-free p–p biradicaloid. *Phys. Chem. Chem. Phys.*, 21(11):5890–5893, 2019. See page 60.

[286] Enrique Mendez-Vega, Mika Maehara, Akshay Hemant Raut, Joel Mieres-Perez, Masashi Tsuge, Yuan-Pern Lee, and Wolfram Sander. Activation of molecular hydrogen by arylcarbenes. *Chem. Eur. J*, 24(70):18801–18808, 2018. See page 60.

[287] Juergen Hinze. Mc-scf. i. the multi-configuration self-consistent-field method. *J. Chem. Phys.*, 59(12):6424–6432, 1973. See page 60.

[288] Per Siegbahn, Anders Heiberg, Björn Roos, and Bernard Levy. A comparison of the super-ci and the newton-raphson scheme in the complete active space scf method. *Phys. Scr*, 21(3-4):323, 1980. See page 60.

[289] Per EM Siegbahn, Jan Almlöf, Anders Heiberg, and Björn O Roos. The complete active space scf (casscf) method in a newton–raphson formulation with application to the hno molecule. *J. Chem. Phys.*, 74(4):2384–2396, 1981. See page 60.

[290] Ernest R Davidson and Charles F Bender. Perturbation theory for multiconfiguration reference states. *Chem. Phys. Let.*, 59(3):369–374, 1978. See page 60.

[291] Bernard Kirtman. Simultaneous calculation of several interacting electronic states by generalized van vleck perturbation theory. *J. Chem. Phys.*, 75(2):798–808, 1981. See page 60.

[292] Kerstin Andersson, Per Aake Malmqvist, Björn O Roos, Andrzej J Sadlej, and Krzysztof Wolinski. Second-order perturbation theory with a casscf reference function. *J. Phys. Chem.*, 94(14):5483–5488, 1990. See page 60.

[293] Walter Kohn. Nobel lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys.*, 71(5):1253, 1999. See page 60.

[294] Soumen Ghosh, Pragya Verma, Christopher J Cramer, Laura Gagliardi, and Donald G Truhlar. Combining wave function methods with density functional theory for excited states. *Chem. Rev.*, 118(15):7249–7292, 2018. See page 60.

[295] Weitao Yang. Generalized adiabatic connection in density functional theory. *J. Chem. Phys.*, 109(23):10107–10110, 1998. See page 60.

[296] R Pollet, F Colonna, T Leininger, H Stoll, H-J Werner, and A Savin. Exchange-correlation energies and correlation holes for some two-and four-electron atoms along a nonlinear adiabatic connection in density functional theory. *Int. J. Quantum Chem.*, 91(2):84–93, 2003. See page 60.

[297] Julien Toulouse, Francois Colonna, and Andreas Savin. Long-range–short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A*, 70(6):062505, 2004. See page 60.

[298] Elisa Rebolini, Julien Toulouse, Andrew M Teale, Trygve Helgaker, and Andreas Savin. Excitation energies along a range-separated adiabatic connection. *J. Chem. Phys.*, 141(4):044123, 2014. See page 60.

[299] János G Angyán, Iann C Gerber, Andreas Savin, and Julien Toulouse. van der waals forces in density functional theory: Perturbational long-range electron-interaction corrections. *Phys. Rev. A*, 72(1):012510, 2005. See page 60.

[300] Erich Goll, Hans-Joachim Werner, and Hermann Stoll. A short-range gradient-corrected density functional in long-range coupled-cluster calculations for rare gas dimers. *Phys. Chem. Chem. Phys.*, 7(23):3917–3923, 2005. See pages 60, 62.

[301] Emmanuel Fromager and Hans Jørgen Aa Jensen. Self-consistent many-body perturbation theory in range-separated density-functional theory: A one-electron reduced-density-matrix-based formulation. *Phys. Rev. A*, 78(2):022504, 2008. See page 60.

[302] Emmanuel Fromager, Renzo Cimiraglia, and Hans Jørgen Aa Jensen. Merging multireference perturbation and density-functional theories by means of range separation: Potential curves for Be_2, Mg_2, and Ca_2. *Phys. Rev. A*, 81(2):024502, 2010. See page 60.

[303] Emmanuel Fromager and Hans Jørgen Aa Jensen. Analysis of self-consistency effects in range-separated density-functional theory with møller-plesset perturbation theory. *J. Chem. Phys.*, 135(3):034116, 2011. See page 60.

[304] David Casanova and Martin Head-Gordon. Restricted active space spin-flip configuration interaction approach: theory, implementation and examples. *Phys. Chem. Chem. Phys.*, 11(42):9779–9790, 2009. See pages 61, 67.

[305] Seymour H Vosko, Leslie Wilk, and Marwan Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of physics*, 58(8):1200–1211, 1980. See pages 62, 92.

[306] Simone Paziani, Saverio Moroni, Paola Gori-Giorgi, and Giovanni B Bachelet. Local-spin-density functional for multideterminant density functional theory. *Physical Review B*, 73(15):155111, 2006. See page 62.

[307] Julien Toulouse, Andreas Savin, and Heinz-Jürgen Flad. Short-range exchange-correlation energy of a uniform electron gas with modified electron–electron interaction. *Int. J. Quantum Chem.*, 100(6):1047–1056, 2004. See pages 62, 67.

[308] Jesper Kielberg Pedersen. *Description of correlation and relativistic effects in calculations of molecular properties.* PhD thesis, University of Southern Denmark, 2004. See page 62.

[309] Martin Fuchs, Y-M Niquet, Xavier Gonze, and Kieron Burke. Describing static correlation in bond dissociation by Kohn–Sham density functional theory. *J. Chem. Phys.*, 122(9):094116, 2005. See page 63.

[310] Walter Heitler and Fritz London. Interaction between neutral atoms and homopolar binding according to quantum mechanics. In *Quantum Chemistry: Classic Scientific Papers*, pages 140–155. World Scientific, 2000. See page 63.

[311] Ángel J Pérez-Jiménez, José M Pérez-Jordá, and Francesc Illas. Density functional theory with alternative spin densities: Application to magnetic systems with localized spins. *J. Chem. Phys.*, 120(1):18–25, 2004. See page 66.

[312] Min-Cheol Kim, Eunji Sim, and Kieron Burke. Understanding and reducing errors in density functional calculations. *Phys. Rev. Lett.*, 111(7):073003, 2013. See page 67.

[313] Lars Hedin. On correlation effects in electron spectroscopies and the gw approximation. *J. Phys.: Condens. Matter*, 11(42):R489, 1999. See page 67.

[314] Seymour H Vosko, Leslie Wilk, and Marwan Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58(8):1200–1211, 1980. See page 67.

[315] Mickaël Hubert, Hans Jørgen Aa Jensen, and Erik D Hedegård. Excitation spectra of nucleobases with multiconfigurational density functional theory. *J. Phys. Chem. A*, 120(1):36–43, 2016. See pages 67, 84.

[316] Yihan Shao, Zhengting Gan, Evgeny Epifanovsky, Andrew TB Gilbert, Michael Wormit, Joerg Kussmann, Adrian W Lange, Andrew Behn, Jia Deng, Xintian Feng, et al. Advances in molecular quantum chemistry contained in the q-chem 4 program package. *Molecular Physics*, 113(2):184–215, 2015. See page 67.

[317] Giovanni Onida, Lucia Reining, and Angel Rubio. Electronic excitations: density-functional versus many-body green's-function approaches. *Rev. Mod. Phys.*, 74:601–659, Jun 2002. See page 70.

[318] Lyudmila V Slipchenko and Anna I Krylov. Singlet-triplet gaps in diradicals by the spin-flip approach: A benchmark study. *J. Chem. Phys.*, 117(10):4694–4708, 2002. See pages 72, 73, 74, 76, 77, and 162.

[319] Yukio Yamaguchi, C David Sherrill, and Henry F Schaefer. The $\tilde{x}^3B_1$, $\tilde{a}^1A_1$, $\tilde{b}_1B_1$, and $\tilde{c}_1A_1$ electronic states of CH_2. *J. Phys. Chem.*, 100(19):7911–7918, 1996. See pages 72, 162.

[320] ST Gibson, JP Greene, and J Berkowitz. Photoionization of the amidogen radical. *J. Chem. Phys.*, 83(9):4319–4328, 1985. See pages 72, 162.

[321] Jeffrey C Stephens, Yukio Yamaguchi, C David Sherrill, and Henry F Schaefer. $\tilde{X}^3B_1, \tilde{a}^1A_1, \tilde{b}^1B_1, \tilde{c}^1\Sigma_g^+$ electronic states of NH_2. *J. Phys. Chem. A*, 102(22):3999–4006, 1998. See pages 72, 162.

[322] C David Sherrill, Matthew L Leininger, Timothy J Van Huis, and Henry F Schaefer III. Structures and vibrational frequencies in the full configuration interaction limit: Predictions for four electronic states of methylene using a triple-zeta plus double polarization (tz2p) basis. *J. Chem. Phys.*, 108(3):1040–1049, 1998. See page 72.

[323] Yukio Yamaguchi, Timothy J Van Huis, C David Sherrill, and Henry F Schaefer III. The $\tilde{x}\ ^1A_1$, $\tilde{a}\ ^3B_1$, $\tilde{a}\ ^1\tilde{b}_1$, and $\tilde{b}^1A_1$ electronic states of SiH_2. *Theor. Chem. Acc.*, 97(1-4):341–349, 1997. See pages 72, 73, and 162.

[324] Timothy J Van Huis, Yukio Yamaguchi, C David Sherrill, and Henry F Schaefer. $\tilde{X}\ ^1a_1$, $\tilde{a}\ ^3b_1$, $\tilde{a}\ ^1\tilde{b}_1$, and $\tilde{a}\ ^1a_1$ electronic states of PH_2^+. *J. Phys. Chem. A*, 101(37):6955–6963, 1997. See pages 72, 73, and 162.

[325] J Berkowitz, JP Greene, H Cho, and B Ruščić. Photoionization mass spectrometric studies of SiH_n (n= 1–4). *J. Chem. Phys.*, 86(3):1235–1248, 1987. See pages 73, 162.

[326] Rafael Escribano and Alain Campargue. Absorption spectroscopy of SiH_2 near 640 nm. *J. Chem. Phys.*, 108(15):6249–6257, 1998. See pages 73, 162.

[327] J Berkowitz and H Cho. A photoionization study of PH: PH_2 revisited. *J. Chem. Phys.*, 90(1):1–6, 1989. See pages 73, 162.

[328] Ana-Maria C Cristian, Yihan Shao, and Anna I Krylov. Bonding patterns in benzene triradicals from structural, spectroscopic, and thermochemical perspectives. *J. Phys. Chem. A*, 108(31):6581–6588, 2004. See page 74.

[329] Paul G Wenthold, Robert R Squires, and WC Lineberger. Ultraviolet photoelectron spectroscopy of the o-, m-, and p-benzyne negative ions. electron affinities and singlet- triplet splittings for o-, m-, and p-benzyne. *J. Am. Chem. Soc.*, 120(21):5279–5290, 1998. See page 76.

[330] Paul Dowd. Trimethylenemethane. *Acc. Chem. Res.*, 5(7):242–248, 1972. See page 76.

[331] Daniel H Ess, Erin R Johnson, Xiangqian Hu, and Weitao Yang. Singlet- triplet energy gaps for diradicals from fractional-spin density-functional theory. *J. Phys. Chem. A*, 115(1):76–83, 2011. See page 76.

[332] Lyudmila V Slipchenko and Anna I Krylov. Electronic structure of the trimethylenemethane diradical in its ground and electronically excited states: Bonding, equilibrium geometries, and vibrational frequencies. *J. Chem. Phys.*, 118(15):6874–6883, 2003. See pages 76, 77, and 78.

[333] Hao Dong, David A Hrovat, Helmut Quast, and Weston Thatcher Borden. Calculations of the relative energies of the low-lying electronic states of 2-methylenedihydrophenalene-1, 3-diyl: Effects of a 1, 8-naphtho bridging group on trimethylenemethane and of a vinylidene bridging group on 1, 8-naphthoquinodimethane. *J. Phys. Chem. A*, 113(5):895–901, 2009. See page 76.

[334] Christopher J Cramer and Bradley A Smith. Trimethylenemethane. comparison of multiconfiguration self-consistent field and density functional methods for a non-kekule hydrocarbon. *J. Phys. Chem.*, 100(23):9664–9670, 1996. See page 76.

[335] David R Yarkony and Henry F Schaefer III. Triplet electronic ground state of trimethylenemethane. *J. Am. Chem. Soc.*, 96(12):3754–3758, 1974. See page 76.

[336] Sheila B Auster, Russell M Pitzer, and Matthew S Platz. Excitation energies in trimethylenemethane derivatives. *J. Am. Chem. Soc.*, 104(14):3812–3815, 1982. See page 76.

[337] Xiangzhu Li and Josef Paldus. Electronic structure of organic diradicals: Evaluation of the performance of coupled-cluster methods. *J. Chem. Phys.*, 129(17):174101, 2008. See page 76.

[338] David Casanova, Lyudmila V Slipchenko, Anna I Krylov, and Martin Head-Gordon. Double spin-flip approach within equation-of-motion coupled cluster and configuration interaction formalisms: Theory, implementation, and examples. *J. Chem. Phys.*, 130(4):044103, 2009. See page 76.

[339] Dorothy S Ginter, Marshall L Ginter, and K Keith Innes. Electronic spectra of the Ga_2, In_2, and Tl_2 molecules. *J. Phys. Chem.*, 69(7):2480–2483, 1965. See page 79.

[340] K Balasubramanian. Spectroscopic constants and potential energy curves of gallium molecules (Ga_2, Ga_2^-, and Ga_2^+). *J. Phys. Chem.*, 94(20):7764–7768, 1990. See page 79.

[341] Kalyan K Das. Ab initio MRD-CI study of the electronic states of the gallium dimer. *J. Phys. B: At. Mol. Opt. Phys.*, 30(4):803, 1997. See page 79.

[342] G Balducci, G Gigli, and Giovanni Meloni. Dissociation energies of the Ga_2, In_2, and GaIn molecules. *J. Chem. Phys.*, 109(11):4384–4388, 1998. See pages 79, 80, and 81.

[343] TK Ghosh, K Tanaka, and Y Mochizuki. Theoretical study of the spectroscopic constants of low-lying states of Ga_2. *J. Mol. Struct. THEOCHEM*, 451(1-2):61–71, 1998. See page 79.

[344] Xiaofeng Tan and Paul J Dagdigian. Electronic spectrum of the gallium dimer. *J. Phys. Chem. A*, 107(15):2642–2649, 2003. See page 79.

[345] Andreas Köhn, Hans-Jörg Himmel, and Benjamin Gaertner. Why Does a Ga_2 Dimer React Spontaneously with H_2, but a Ga Atom Does Not?—A Detailed Quantum Chemical Investigation of the Differences in Reactivity Between Ga Atoms and Ga_2 Dimers, in Combination with Experimental Results. *Chem. Eur. J*, 9(16):3909–3919, 2003. See pages 79, 81.

[346] N Gaston and AJ Parker. On the bonding of Ga_2, structures of Ga_n clusters and the relation to the bulk structure of gallium. *Chem. Phys. Let.*, 501(4-6):375–378, 2011. See pages 79, 81.

[347] Ralf Tonner and Nicola Gaston. The dimeric nature of bonding in gallium: from small clusters to the α-gallium phase. *Phys. Chem. Chem. Phys.*, 16(44):24244–24249, 2014. See page 79.

[348] Nedko Drebov, Florian Weigend, and Reinhart Ahlrichs. Structures and properties of neutral gallium clusters: A theoretical investigation. *J. Chem. Phys.*, 135(4):044314, 2011. See pages 79, 80, and 81.

[349] José Aarón Rodríguez-Jiménez, Erik Díaz-Cervantes, Faustino Aguilera-Granja, and Juvencio Robles. Computational study of Ga_nAs_m (m + n= 2–9) clusters using dft calculations. *J. Nanopart. Res.*, 21(11):219, 2019. See page 79.

[350] ZL Xiao, RH Hauge, and JL Margrave. Cryogenic reactions of gallium with molecular hydrogen and methane. *Inorg. Chem.*, 32(5):642–646, 1993. See page 79.

[351] Hans-Jörg Himmel, Laurent Manceron, Anthony J Downs, and Pluton Pullumbi. Formation and characterization of the gallium and indium subhydride molecules Ga_2H_2 and In_2H_2: a matrix isolation study. *J. Am. Chem. Soc.*, 124(16):4448–4457, 2002. See page 79.

[352] P Pullumbi, C Mijoule, L Manceron, and Y Bouteiller. Aluminium, gallium and indium dihydrides. an ir matrix isolation and ab initio study. *Chem. Phys.*, 185(1):13–24, 1994. See page 79.

[353] Lon B Knight Jr, John J Banisaukas III, Robert Babb, and Ernest R Davidson. Electron spin resonance matrix isolation and ab initio theoretical investigations of $^{69,71}GaH_2$, $^{69,71}GaD_2$, $H^{69,71}GaCH_3$, and $D^{69,71}GaCD_3$. *J. Chem. Phys.*, 105(16):6607–6615, 1996. See page 79.

[354] Florian Weigend and Reinhart Ahlrichs. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for h to rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.*, 7(18):3297–3305, 2005. See page 79.

[355] Yi Zhao, Wenguo Xu, Qianshu Li, Yaoming Xie, and Henry F Schaefer. Gallium clusters Ga_n (n= 1- 6): Structures, thermochemistry, and electron affinities. *J. Phys. Chem. A*, 108(36):7448–7459, 2004. See pages 80, 81.

[356] Björn O Roos, Roland Lindh, Per-Åke Malmqvist, Valera Veryazov, and Per-Olof Widmark. Main group atoms and dimers studied with a new relativistic ano basis set. *J. Phys. Chem. A*, 108(15):2851–2858, 2004. See pages 80, 81.

[357] Dennis P Carroll, Harris J Silverstone, and Robert Melville Metzger. Piecewise polynomial configuration interaction natural orbital study of 1 s 2 helium. *The Journal of Chemical Physics*, 71(10):4142–4163, 1979. See page 84.

[358] Pablo Echenique and José Luis Alonso. A mathematical and computational review of hartree–fock scf methods in quantum chemistry. *Molecular Physics*, 105(23-24):3057–3098, 2007. See page 84.

[359] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140:A1133–A1138, Nov 1965. See page 84.

[360] Hedegård Erik Donovan. *Multi-Configurational Density Functional Theory: Progress and Challenges*, chapter 3, pages 47–75. John Wiley & Sons, Ltd, 2020. See page 84.

[361] Dayou Zhang, Matthew R Hermes, Laura Gagliardi, and Donald G Truhlar. Multiconfiguration density-coherence functional theory. *Journal of Chemical Theory and Computation*, 17(5):2775–2782, 2021. See page 84.

[362] Hans Jo/rgen Aa Jensen, Poul Jo/rgensen, Hans Ågren, and Jeppe Olsen. Second-order mo/ller–plesset perturbation theory as a configuration and orbital generator in multiconfiguration self-consistent field calculations. *The Journal of Chemical Physics*, 88(6):3834–3839, 1988. See page 84.

[363] Christopher J Stein and Markus Reiher. Automated selection of active orbital spaces. *Journal of Chemical Theory and Computation*, 12(4):1760–1771, 2016. See page 84.

[364] Abhishek Khedkar and Michael Roemelt. Active space selection based on natural orbital occupation numbers from n-electron valence perturbation theory. *Journal of Chemical Theory and Computation*, 15(6):3522–3536, 2019. See page 84.

[365] Detlef Doehnert and Jaroslav Koutecky. Occupation numbers of natural orbitals as a criterion for biradical character. different kinds of biradicals. *Journal of the American Chemical Society*, 102(6):1789–1796, 1980. See page 84.

[366] Tullimilli Y Gopalakrishna, Wangdong Zeng, Xuefeng Lu, and Jishan Wu. From open-shell singlet diradicaloids to polyradicaloids. *Chemical Communications*, 54(18):2186–2199, 2018. See page 84.

[367] Thijs Stuyver, Bo Chen, Tao Zeng, Paul Geerlings, Frank De Proft, and Roald Hoffmann. Do diradicals behave like radicals? *Chemical Reviews*, 119(21):11291–11351, 2019. See page 84.

[368] C. Gutlé and A. Savin. Orbital spaces and density-functional theory. *Physical Review A*, 75:032519, Mar 2007. See page 84.

[369] Elliott H Lieb. Density functionals for coulomb systems. *Inequalities: Selecta of Elliott H. Lieb*, pages 269–303, 2002. See page 85.

[370] Aron J. Cohen, Paula Mori-Sánchez, and Weitao Yang. Insights into current limitations of density functional theory. *Science*, 321(5890):792–794, 2008. See page 86.

[371] Jürgen Gräfenstein and Dieter Cremer. Development of a cas-dft method covering non-dynamical and dynamical electron correlation in a balanced way. *Molecular Physics*, 103(2-3):279–308, 2005. See page 86.

[372] Markus Holzmann, Bernard Bernu, Carlo Pierleoni, Jeremy McMinis, David M. Ceperley, Valerio Olevano, and Luigi Delle Site. Momentum distribution of the homogeneous electron gas. *Physical Review Letters*, 107:110402, Sep 2011. See page 86.

[373] Equivalently to fixing k_c, the threshold may be imposed by defining the value for the integral (from k_c to ∞) over the momentum distribution. See page 88.

[374] Paola Gori-Giorgi and Paul Ziesche. Momentum distribution of the uniform electron gas: Improved parametrization and exact limits of the cumulant expansion. *Physical Review B*, 66(23):235116, 2002. See page 88.

[375] JC Kimball. Short-range correlations and the structure factor and momentum distribution of electrons. *Journal of Physics A: Mathematical and General*, 8(9):1513, 1975. See page 88.

[376] H Yasuhara and Y Kawazoe. A note on the momentum distribution function for an electron gas. *Physica A: Statistical Mechanics and its Applications*, 85(2):416–424, 1976. See page 88.

[377] Paola Gori-Giorgi and John P Perdew. Short-range correlation in the uniform electron gas: Extended overhauser model. *Physical Review B*, 64(15):155102, 2001. See page 88.

[378] E Paijanne and J Arponen. Electron liquid in the collective description. vi. static properties of the homogeneous electron gas. *Journal of Physics C: Solid State Physics*, 15(12):2683, 1982. See page 89.

[379] David L Freeman. Coupled-cluster expansion applied to the electron gas: Inclusion of ring and exchange effects. *Physical Review B*, 15(12):5512, 1977. See page 89.

[380] D. Casanova and M. Head-Gordon. Restricted active space spin-flip configuration interaction approach: Theory, implementation and examples. *Physical Chemistry Chemical Physics*, 11:9779–9790, 2009. See pages 89, 134.

[381] Evgeny Epifanovsky, Andrew T. B. Gilbert, Xintian Feng, Joonho Lee, Yuezhi Mao, Narbe Mardirossian, Pavel Pokhilko, Alec F. White, Marc P. Coons, Adrian L. Dempwolff, Zhengting Gan, Diptarka Hait, Paul R. Horn, Leif D. Jacobson, Ilya Kaliman, Jörg Kussmann, Adrian W. Lange, Ka Un Lao, Daniel S. Levine, Jie Liu, Simon C. McKenzie, Adrian F. Morrison, Kaushik D. Nanda, Felix Plasser, Dirk R. Rehn, Marta L. Vidal, Zhi-Qiang You, Ying Zhu, Bushra Alam, Benjamin J. Albrecht, Abdulrahman Aldossary, Ethan Alguire, Josefine H. Andersen, Vishikh Athavale, Dennis Barton, Khadiza Begam, Andrew Behn, Nicole Bellonzi, Yves A. Bernard, Eric J. Berquist, Hugh G. A. Burton, Abel Carreras, Kevin Carter-Fenk, Romit Chakraborty, Alan D. Chien, Kristina D. Closser, Vale Cofer-Shabica, Saswata Dasgupta, Marc de Wergifosse, Jia Deng, Michael Diedenhofen, Hainam Do, Sebastian Ehlert, Po-Tung Fang, Shervin Fatehi, Qingguo Feng, Triet Friedhoff, James Gayvert, Qinghui Ge, Gergely Gidofalvi, Matthew Goldey, Joe Gomes, Cristina E. González-Espinoza, Sahil Gulania, Anastasia O. Gunina, Magnus W. D. Hanson-Heine, Phillip H. P. Harbach, Andreas Hauser, Michael F. Herbst, Mario Hernández Vera, Manuel Hodecker, Zachary C. Holden, Shannon Houck, Xunkun Huang, Kerwin Hui, Bang C. Huynh, Maxim Ivanov, Ádám Jász, Hyunjun Ji, Hanjie Jiang, Benjamin Kaduk, Sven Kähler, Kirill Khistyaev, Jaehoon Kim, Gergely Kis, Phil Klunzinger, Zsuzsanna Koczor-Benda, Joong Hoon Koh, Dimitri Kosenkov, Laura Koulias, Tim Kowalczyk, Caroline M. Krauter, Karl Kue, Alexander Kunitsa, Thomas Kus, István Ladjánszki, Arie Landau, Keith V. Lawler, Daniel Lefrancois, Susi Lehtola, Run R. Li, Yi-Pei Li, Jiashu Liang, Marcus Liebenthal, Hung-Hsuan Lin, You-Sheng Lin, Fenglai Liu, Kuan-Yu Liu, Matthias Loipersberger, Arne Luenser, Aaditya Manjanath, Prashant Manohar, Erum Mansoor, Sam F. Manzer, Shan-Ping Mao, Aleksandr V. Marenich, Thomas Markovich, Stephen Mason, Simon A. Maurer, Peter F. McLaughlin, Maximilian F. S. J. Menger, Jan-Michael Mewes, Stefanie A. Mewes, Pierpaolo Morgante, J. Wayne Mullinax, Katherine J. Oosterbaan, Garrette Paran, Alexander C. Paul, Suranjan K. Paul, Fabijan Pavošević, Zheng Pei, Stefan Prager, Emil I. Proynov, Ádám Rák, Eloy Ramos-Cordoba, Bhaskar Rana, Alan E. Rask, Adam Rettig, Ryan M. Richard, Fazle Rob, Elliot Rossomme, Tarek Scheele, Maximilian Scheurer, Matthias Schneider, Nickolai Sergueev, Shaama M. Sharada, Wojciech Skomorowski, David W. Small, Christopher J. Stein, Yu-Chuan Su, Eric J. Sundstrom, Zhen Tao, Jonathan Thirman, Gábor J. Tornai, Takashi Tsuchimochi, Norm M. Tubman, Srimukh Prasad Veccham, Oleg Vydrov, Jan Wenzel, Jon Witte, Atsushi

Yamada, Kun Yao, Sina Yeganeh, Shane R. Yost, Alexander Zech, Igor Ying Zhang, Xing Zhang, Yu Zhang, Dmitry Zuev, Alán Aspuru-Guzik, Alexis T. Bell, Nicholas A. Besley, Ksenia B. Bravaya, Bernard R. Brooks, David Casanova, Jeng-Da Chai, Sonia Coriani, Christopher J. Cramer, György Cserey, A. Eugene DePrince, Robert A. DiStasio, Andreas Dreuw, Barry D. Dunietz, Thomas R. Furlani, William A. Goddard, Sharon Hammes-Schiffer, Teresa Head-Gordon, Warren J. Hehre, Chao-Ping Hsu, Thomas-C. Jagau, Yousung Jung, Andreas Klamt, Jing Kong, Daniel S. Lambrecht, WanZhen Liang, Nicholas J. Mayhall, C. William McCurdy, Jeffrey B. Neaton, Christian Ochsenfeld, John A. Parkhill, Roberto Peverati, Vitaly A. Rassolov, Yihan Shao, Lyudmila V. Slipchenko, Tim Stauch, Ryan P. Steele, Joseph E. Subotnik, Alex J. W. Thom, Alexandre Tkatchenko, Donald G. Truhlar, Troy Van Voorhis, Tomasz A. Wesolowski, K. Birgitta Whaley, H. Lee Woodcock, Paul M. Zimmerman, Shirin Faraji, Peter M. W. Gill, Martin Head-Gordon, John M. Herbert, and Anna I. Krylov. Software for the frontiers of quantum chemistry: An overview of developments in the q-chem 5 package. *The Journal of Physical Chemistry*, 155(8):084801, 2021. See pages 92, 136.

[382] John F Dobson, Giovanni Vignale, and Mukunda P Das. Electronic density functional theory: recent progress and new directions. 2013. See page 93.

[383] Andreas Savin, François Colonna, and Jean-Marie Teuler. Adiabatic coupling in the helium and the beryllium series. *Electronic Density Functional Theory: Recent Progress and New Directions*, pages 69–80, 1998. See page 93.

[384] John F Dobson and Mukunda P Das. Brief introduction to density functional theory. In *Electronic Density Functional Theory: Recent Progress and New Directions*, pages 3–18. Springer, 1998. See page 93.

[385] JP Perdew, ER McMullen, and Alex Zunger. Density-functional theory of the correlation energy in atoms and ions: a simple analytic model and a challenge. *Physical Review A*, 23(6):2785, 1981. See page 93.

[386] Even the FCI treatment fails for an accurate quantitative estimation of the correlation energy. Errors at the FCI level with respect to exact values arise due to the limited size of the basis sets, cusp condition violation and to the extent of linear dependencies among the primitives. FCI correlation energies can be improved by following a careful optimization of the parameters in the basis set (see Refs. [490, 491]). See page 93.

[387] Walter Heitler and Fritz London. Interaction between neutral atoms and homopolar binding according to quantum mechanics. In *Quantum Chemistry: Classic Scientific Papers*, pages 140–155. World Scientific, 2000. See page 95.

[388] Martin Fuchs, Y-M Niquet, Xavier Gonze, and Kieron Burke. Describing static correlation in bond dissociation by kohn–sham density functional theory. *The Journal of Chemical Physics*, 122(9), 2005. See page 95.

[389] David Feller, Kirk A Peterson, and Ernest R Davidson. A systematic approach to vertically excited states of ethylene using configuration interaction and coupled cluster techniques. *The Journal of Chemical Physics*, 141(10):104302, 2014. See pages 96, 141.

[390] Arthur F Voter, Marvin M Goodgame, and William A Goddard III. Intraatomic exchange and the violation of hund's rule in twisted ethylene. *Chemical Physics*, 98(1):7–14, 1985. See pages 96, 141.

[391] Anna I Krylov, C David Sherrill, Edward FC Byrd, and Martin Head-Gordon. Size-consistent wave functions for nondynamical correlation energy: The valence active space optimized orbital coupled-cluster doubles model. *The Journal of Chemical Physics*, 109(24):10669–10678, 1998. See pages 96, 97.

[392] Anna I Krylov. Spin-contamination of coupled-cluster wave functions. *The Journal of Chemical Physics*, 113(15):6052–6062, 2000. See page 96.

[393] David Casanova and Martin Head-Gordon. The spin-flip extended single excitation configuration interaction method. *The Journal of Chemical Physics*, 129(6), 2008. See page 96.

[394] Yihan Shao, Martin Head-Gordon, and Anna I Krylov. The spin–flip approach within time-dependent density functional theory: Theory and applications to diradicals. *The Journal of Chemical Physics*, 118(11):4807–4818, 2003. See page 96.

[395] Henryk A Witek, Haruyuki Nakano, and Kimihiko Hirao. Multireference perturbation theory with optimized partitioning. ii. applications to molecular systems. *Journal of Computational Chemistry*, 24(12):1390–1400, 2003. See page 96.

[396] Emilio San-Fabián and Federico Moscardó. Cyclobutadiene automerization and rotation of ethylene: Energetics of the barriers by using spin-polarized wave functions. *Journal of Computational Chemistry*, 35(18):1356–1363, 2014. See pages 96, 97.

[397] AJ Merer and Robert S Mulliken. Ultraviolet spectra and excited states of ethylene and its alkyl derivatives. *Chemical Reviews*, 69(5):639–656, 1969. See pages 96, 97.

[398] Bernhard Gemein and Sigrid D Peyerimhoff. Radiationless transitions between the first excited triplet state and the singlet ground state in ethylene: A theoretical study. *The Journal of Physical Chemistry*, 100(50):19257–19267, 1996. See pages 96, 97.

[399] David Danovich, Christel M Marian, Thomas Neuheuser, Sigrid D Peyerimhoff, and Sason Shaik. Spin- Orbit Coupling Patterns Induced by Twist and Pyramidalization Modes in C_2H_4: A Quantitative Study and a Qualitative Analysis. *The Journal of Physical Chemistry A*, 102(29):5923–5936, 1998. See pages 96, 141, 142, and 150.

[400] Jan AlmlÖf, Bradley J Deleeuw, Peter R Taylor, Charles W Bauschlicher Jr, and Per Siegbahn. The dissociation energy of n2. *International Journal of Quantum Chemistry*, 36(S23):345–354, 1989. See pages 97, 98, 117, and 118.

[401] Helena Larsen, Jeppe Olsen, Poul Jørgensen, and Ove Christiansen. Full configuration interaction benchmarking of coupled-cluster models for the lowest singlet energy surfaces of n 2. *The Journal of Chemical Physics*, 113(16):6677–6686, 2000. See page 97.

[402] Stephen R Langhoff, Charles W Bauschlicher Jr, and Peter R Taylor. Accurate ab initio calculations for the ground states of n2, o2 and f2. *Chemical physics letters*, 135(6):543–548, 1987. See pages 97, 117.

[403] Reinhart Ahlrichs, Peter Scharf, and Karol Jankowski. The impact of higher polarization basis functions on molecular ab initio results ii. the ground states of co, n2, o2, and f2. *Chemical physics*, 98(3):381–386, 1985. See pages 97, 117.

[404] Hans-Joachim Werner and Peter J Knowles. Accurate multireference configuration interaction calculations of the potential energy function and the dissociation energy of n2. *The Journal of chemical physics*, 94(2):1264–1270, 1991. See page 97.

[405] Oliver Marchetti and Hans-Joachim Werner. Accurate calculations of intermolecular interaction energies using explicitly correlated wave functions. *Physical Chemistry Chemical Physics*, 10(23):3400–3409, 2008. See page 102.

[406] Toru Shiozaki and Hans-Joachim Werner. Explicitly correlated multireference configuration interaction with multiple reference functions: Avoided crossings and conical intersections. *The Journal of chemical physics*, 134(18), 2011. See page 102.

[407] Toru Shiozaki and Hans-Joachim Werner. Multireference explicitly correlated f12 theories. *Molecular Physics*, 111(5):607–630, 2013. See page 102.

[408] Samuel Francis Boys and Nicholas Charles Handy. A first solution, for lih, of a molecular transcorrelated wave equation by means of restricted numerical integration. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 311(1505):309–329, 1969. See page 102.

[409] Samuel Francis Boys and Nicholas Charles Handy. A condition to remove the indeterminacy in interelectronic correlation functions. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 309(1497):209–220, 1969. See page 102.

[410] Samuel Francis Boys and Nicholas Charles Handy. A calculation for the energies and wavefunctions for states of neon with full electronic correlation accuracy. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 310(1500):63–78, 1969. See page 102.

[411] NC Handy. Energies and expectation values for be by the transcorrelated method. *The Journal of Chemical Physics*, 51(8):3205–3212, 1969. See pages 102, 107.

[412] Renato Colle and Oriano Salvetti. Approximate calculation of the correlation energy for the closed shells. *Theoretica chimica acta*, 37:329–334, 1975. See pages 103, 104.

[413] A John Coleman. Structure of fermion density matrices. *Reviews of modern Physics*, 35(3):668, 1963. See page 103.

[414] WA Bingel. A physical interpretation of the cusp conditions for molecular wave functions. *Theoretica chimica acta*, 8(1):54–61, 1967. See page 104.

[415] Eugene Wigner. On the interaction of electrons in metals. *Physical Review*, 46(11):1002, 1934. See pages 104, 166.

[416] Enrico Clementi. Correlation energy for atomic systems. *The Journal of Chemical Physics*, 38(9):2248–2256, 1963. See page 106.

[417] Enrico Clementi. Correlation energy for atomic systems. ii. isoelectronic series with 11 to 18 electrons. *The Journal of Chemical Physics*, 39(1):175–179, 1963. See page 106.

[418] Michel Dupuis, John Rys, and Harry F King. Evaluation of molecular integrals over gaussian basis functions. *The Journal of Chemical Physics*, 65(1):111–116, 1976. See pages 107, 193.

[419] Michael Böhme. A fortran library to numerically calculate the boys function. `https://github.com/micb25/libboys`, 2022. See page 107.

[420] Qiming Sun, Timothy C Berkelbach, Nick S Blunt, George H Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D McClain, Elvira R Sayfutyarova, Sandeep Sharma, et al. Pyscf: the python-based simulations of chemistry framework. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 8(1):e1340, 2018. See page 108.

[421] Thom H Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *The Journal of chemical physics*, 90(2):1007–1023, 1989. See page 108.

[422] Richard FW Bader, Andrew Streitwieser, Arnst Neuhaus, Keith E Laidig, and Peter Speers. Electron delocalization and the fermi hole. *Journal of the American Chemical Society*, 118(21):4959–4965, 1996. See page 108.

[423] Caio L Firme, Octávio Augusto Ceva Antunes, and Pierre M Esteves. Relation between bond order and delocalization index of qtaim. *Chemical Physics Letters*, 468(4-6):129–133, 2009. See page 108.

[424] E. Matito and F. Feixas. Dmn program. Universitat de Girona (Spain) and Uniwersytet Szczeciński (Poland), 2009. See page 108.

[425] M. Rodríguez-Mayorga. Rho2-ops: 2-dm opertions. Institut de Química Computacional i Catàlisi (IQCC), Universitat de Girona (spain)., 2022. See page 108.

[426] Jerzy Cioslowski and Guanghua Liu. Fast evaluation of electron intracule and extracule densities on large grids of points. *The Journal of chemical physics*, 105(10):4151–4158, 1996. See page 108.

[427] Richard FW Bader. Atoms in molecules. *Accounts of chemical research*, 18(1):9–15, 1985. See page 108.

[428] P Shyam Vinod Kumar, V Raghavendra, and V Subramanian. Bader's theory of atoms in molecules (aim) and its applications to chemical bonding. *Journal of Chemical Sciences*, 128:1527–1536, 2016. See page 108.

[429] Todd A. Keith. Aimall (version 19.10.12). TK Gristmill Software, Overland Park KS, USA, 2019. See page 108.

[430] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian˜16 Revision C.01, 2016. Gaussian Inc. Wallingford CT. See page 108.

[431] Charlotte Emma Moore. *Atomic energy levels as derived from the analyses of optical spectra*, volume 1. US Department of Commerce, National Bureau of Standards, 1949. See page 111.

[432] David Robert Bates. *Atomic and molecular processes*, volume 13. Elsevier, 2012. See page 111.

[433] Moa K Kristiansson, Kiattichart Chartkunchand, Gustav Eklund, Odd M Hole, Emma K Anderson, Nathalie de Ruette, Magdalena Kamińska, Najeeb Punnakayathil, José E Navarro-Navarrete, Stefan Sigurdsson, et al. High-precision electron affinity of oxygen. *Nature Communications*, 13(1):5906, 2022. See page 111.

[434] Jochen Autschbach. Perspective: relativistic effects. *The Journal of Chemical Physics*, 136(15):150902, 2012. See page 133.

[435] Christel M Marian. Spin–orbit coupling and intersystem crossing in molecules. *WIREs Computational Molecular Science*, 2(2):187–203, 2012. See page 133.

[436] Michael Kasha. Phosphorescence and the Role of the Triplet State in the Electronic Excitation of Complex Molecules. *Chemical Reviews*, 41(2):401–419, 1947. See page 133.

[437] Andrew J. Jenkins, Hang Hu, Lixin Lu, Michael J. Frisch, and Xiaosong Li. Two-component multireference restricted active space configuration interaction for the computation of l-edge x-ray absorption spectra. *Journal of Chemical Theory and Computation*, 18(1):141–150, 2022. PMID: 34908414. See page 133.

[438] David R Yarkony. Spin-forbidden chemistry within the breit-pauli approximation. *International Reviews in Physical Chemistry*, 11(2):195–242, 1992. See page 133.

[439] Martin Klessinger and Josef Michl. *Excited states and photochemistry of organic molecules*. VCH, New York, 1995. See pages 133, 134.

[440] RJ Cvetanović. Addition of atoms to olefins in the gas phase. *Advances in Photochemistry*, 1:115–182, 1963. See page 134.

[441] Hassan Sabbah, Ludovic Biennier, Ian R Sims, Yuri Georgievskii, Stephen J Klippenstein, and Ian WM Smith. Understanding reactivity at very low temperatures: The reactions of oxygen atoms with alkenes. *Science*, 317(5834):102–105, 2007. See page 134.

[442] DL Singleton, S Furuyama, RJ Cvetanović, and RS Irwin. Temperature dependence of the rate constants for the reactions o (^{3}p)+ 2, 3-dimethyl-2-butene and o (^{3}p) + no + m determined by a phase shift technique. *The Journal of Chemical Physics*, 63(2):1003–1007, 1975. See page 134.

[443] Craig A Taatjes, David L Osborn, Talitha M Selby, Giovanni Meloni, Adam J Trevitt, Evgeny Epifanovsky, Anna I Krylov, Baptiste Sirjean, Enoch Dames, and Hai Wang. Products of the benzene + o (^{3}p) reaction. *The Journal of Physical Chemistry A*, 114(9):3355–3370, 2010. See page 134.

[444] M Dupuis, JJ Wendoloski, T Takada, and W_ A Lester Jr. Theoretical study of electrophilic addition: O (^{3}p) + c_2h_4. *The Journal of Chemical Physics*, 76(1):481–487, 1982. See page 134.

[445] Shaolei Zhao, Weiqiang Wu, Hongmei Zhao, Huan Wang, Chunfan Yang, Kunhui Liu, and Hongmei Su. Adiabatic and nonadiabatic reaction pathways of the o (^{3}p) with propyne. *The Journal of Physical Chemistry A*, 113(1):23–34, 2009. See page 134.

[446] Thanh Lam Nguyen, Jozef Peeters, and Luc Vereecken. Theoretical reinvestigation of the o (3p)+ c6h6 reaction: Quantum chemical and statistical rate calculations. *The Journal of Physical Chemistry A*, 111(19):3836–3849, 2007. See page 134.

[447] Bina Fu, Yong-Chang Han, Joel M Bowman, Luca Angelucci, Nadia Balucani, Francesca Leonori, and Piergiorgio Casavecchia. Intersystem crossing and dynamics in o (3p)+ c2h4 multichannel reaction: Experiment validates theory. *Proceedings of the National Academy of Sciences*, 109(25):9733–9738, 2012. See page 134.

[448] Thanh Lam Nguyen, Luc Vereecken, Xin Juan Hou, Minh Tho Nguyen, and Jozef Peeters. Potential energy surfaces, product distributions and thermal rate coefficients of the reaction of o (^{3}p) with c_2h_4 (xag): a comprehensive theoretical study. *The Journal of Physical Chemistry A*, 109(33):7489–7499, 2005. See page 134.

[449] Devin Hodgson, Hai-Yue Zhang, Mark R Nimlos, and J Thomas McKinnon. Quantum chemical and rrkm investigation of the elementary channels of the reaction $c_6h_6^+$ o (^{3}p). *The Journal of Physical Chemistry A*, 105(17):4316–4327, 2001. See page 134.

[450] Alejandro Avilés, Ana E Torres, and Fernando Colmenares. Low-energy pathways found for the nh_3 activation and h_2 elimination by the werner-type complexes $m(nh_3)^{4+}$ (m= fe, ru and os). *ChemistrySelect*, 5(47):15039–15048, 2020. See page 134.

[451] Nicholas J Turro and Alain Devaquet. Chemiexcitation mechanims. role of symmetry and spin-orbit coupling in diradicals. *Journal of the American Chemical Society*, 97(13):3859–3862, 1975. See page 134.

[452] Attila Tajti, Péter G. Szalay, Attila G. Császár, Mihály Kállay, Jürgen Gauss, Edward F. Valeev, Bradley A. Flowers, Juana Vázquez, and John F. Stanton. Heat: High accuracy extrapolated ab initio thermochemistry. *The Journal of Chemical Physics*, 121(23):11599–11613, 2004. See page 134.

[453] Hans A Bethe and Edwin E Salpeter. *Quantum mechanics of one-and two-electron atoms*. Springer Science & Business Media, 2012. See page 134.

[454] Bernd A Heß, Christel M Marian, Ulf Wahlgren, and Odd Gropen. A mean-field spin-orbit method applicable to correlated wavefunctions. *Chemical Physics Letters*, 251(5-6):365–371, 1996. See page 135.

[455] Evgeny Epifanovsky, Kerstin Klein, Stella Stopkowicz, Jürgen Gauss, and Anna I Krylov. Spin-orbit couplings within the equation-of-motion coupled-cluster framework: Theory, implementation, and benchmark calculations. *The Journal of chemical physics*, 143(6):064102, 2015. See page 135.

[456] Pavel Pokhilko, Evgeny Epifanovsky, and Anna I Krylov. General framework for calculating spin–orbit couplings using spinless one-particle density matrices: Theory and application to the equation-of-motion coupled-cluster wave functions. *The Journal of chemical physics*, 151(3):034106, 2019. See pages 135, 223.

[457] Dmitri G Fedorov and Mark S Gordon. A study of the relative importance of one and two-electron contributions to spin–orbit coupling. *The Journal of Chemical Physics*, 112(13):5611–5623, 2000. See pages 136, 142.

[458] Dmitri G Fedorov, Shiro Koseki, Michael W Schmidt, and Mark S Gordon. Spin-orbit coupling in molecules: chemistry beyond the adiabatic approximation. *International Reviews in Physical Chemistry*, 22(3):551–592, 2003. See pages 136, 142.

[459] CEH Bawn and RF Hunter. The trimethylene biradical. *Transactions of the Faraday Society*, 34:608–613, 1938. See page 137.

[460] Paul B Condit and Robert G Bergman. On the nature of diradical intermediates formed during the thermal and photochemical decomposition of bicyclic azo-compounds. *Journal of the Chemical Society, Chemical Communications*, (1):4b–6, 1971. See page 137.

[461] PJ Hay, WJ Hunt, and WA Goddard III. Theoretical investigations of the trimethylene biradical. *Journal of the American Chemical Society*, 94(2):638–640, 1972. See page 137.

[462] Charles Doubleday Jr, James W McIver Jr, and Michael Page. Chemistry of triplet vs. singlet biradicals. An ab initio MCSCF study of trimethylene. *Journal of the American Chemical Society*, 104(24):6533–6542, 1982. See pages 137, 139.

[463] Yukio Yamaguchi and Henry F Schaefer III. Equilibrium geometries for triplet trimethylene, $\cdot CH_2CH_2CH_2\cdot$. *Journal of the American Chemical Society*, 106(18):5115–5118, 1984. See page 137.

[464] Thomas R Furlani and Harry F King. Theory of spin-orbit coupling. Application to singlet–triplet interaction in the trimethylene biradical. *The Journal of Chemical Physics*, 82(12):5577–5583, 1985. See pages 137, 139.

[465] Marcus Böckmann, Martin Klessinger, and Michael C Zerner. Spin- orbit coupling in organic molecules: A semiempirical configuration interaction approach toward triplet state reactivity. *The Journal of Physical Chemistry*, 100(25):10570–10579, 1996. See page 137.

[466] EvelynáM Goldfield. Quantum dynamics of the trimethylene biradical Stereomutation of cyclopropane and unimolecular decay. *Faraday Discussions*, 110:185–205, 1998. See page 137.

[467] John A Horsley, Yves Jean, Carl Moser, Lionel Salem, Richard M Stevens, and James S Wright. Organic transition state. *Journal of the American Chemical Society*, 94(1):279–282, 1972. See page 137.

[468] Louis Carlacci, Charles Doubleday Jr, Thomas R Furlani, Harry F King, and James W McIver Jr. Spin-orbit coupling in biradicals. Ab initio MCSCF calculations on trimethylene and the methyl-methyl radical pair. *Journal of the American Chemical Society*, 109(18):5323–5329, 1987. See page 137.

[469] Mostafa A. El-Sayed. Triplet state. its radiative and nonradiative properties. *Accounts of Chemical Research*, 1(1):8–16, 01 1968. See page 138.

[470] A. I. Krylov, C. D. Sherrill, E. F. C. Byrd, and M. Head-Gordon. Size-consistent wavefunctions for non-dynamical correlation energy: The valence active space optimized orbital coupled-cluster doubles model. *The Journal of Chemical Physics*, 109(24):10669–10678, 1998. See page 141.

[471] A. I. Krylov. Spin-contamination in coupled cluster wavefunctions. *The Journal of Chemical Physics*, 113(15):6052–6062, 2000. See page 141.

[472] D. Casanova and M. Head-Gordon. The spin-flip extended single excitation configuration interaction method. *The Journal of Chemical Physics*, 129:064104, 2008. See page 141.

[473] Josef Michl. Spin-orbit coupling in biradicals. 1. the 2-electrons-in-2-orbitals model revisited. *Journal of the American Chemical Society*, 118(15):3568–3579, 1996. See page 141.

[474] Michael Filatov, Todd J Martínez, and Kwang S Kim. Description of ground and excited electronic states by ensemble density functional method with extended active space. *The Journal of Chemical Physics*, 147(6):064104, 2017. See page 143.

[475] William R Ware. Oxygen quenching of fluorescence in solution: an experimental study of the diffusion process. *The Journal of Physical Chemistry*, 66(3):455–458, 1962. See page 150.

[476] Isadore Berlman. *Handbook of fluorescence spectra of aromatic molecules*. Elsevier, New York, 2012. See page 150.

[477] Samuel Lenher. The reaction between oxygen and ethylene. I. *Journal of the American Chemical Society*, 53(10):3737–3751, 1931. See page 150.

[478] Samuel Lenher. The reaction between oxygen and ethylene. II. *Journal of the American Chemical Society*, 53(10):3752–3765, 1931. See page 150.

[479] Ronald George Wreyford Norrish and K Porter. Some features of the gas phase oxidation of n-butenes. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 272(1349):164–191, 1963. See page 150.

[480] Robert E Burk, Bruce G Baldwin, and CH Whitacre. Thermal reactions of ethylene. *Industrial & Engineering Chemistry*, 29(3):326–330, 1937. See page 150.

[481] John Duncan Mullen and G Skirrow. Gas-phase oxidation of propylene. *Industrial & Engineering Chemistry Research*, 244(1238):312–330, 1958. See page 150.

[482] AJ Harding and Ronald George Wreyford Norrish. The role of formaldehyde in the oxidation of ethylene. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 212(1110):291–311, 1952. See page 150.

[483] Satoshi Inagaki, Shinichi Yamabe, Hiroshi Fujimoto, and Kenichi Fukui. A consideration of orbital interaction in the reaction of $^1\delta_g$ molecular oxygen with ethylene derivatives. *Bulletin of the Chemical Society of Japan*, 45(12):3510–3514, 1972. See page 150.

[484] Chia-Chung Chen and Marye Anne Fox. Reaction path calculations for the interaction of the ethylene radical cation with triplet oxygen. *Journal of Computational Chemistry*, 4(4):488–493, 1983. See page 150.

[485] Frederik Thorning, Frank Jensen, and Peter R. Ogilby. The oxygen–organic molecule photosystem: revisiting the past, recalibrating the present, and redefining the future. *Photochemical & Photobiological Sciences*, 21(7):1133–1141, 2022. See page 150.

[486] Frederik Thorning, Frank Jensen, and Peter R. Ogilby. Geometry dependence of spin–orbit coupling in complexes of molecular oxygen with atoms, h_2, or organic molecules. *The Journal of Physical Chemistry A*, 126(6):834–844, 2022. See pages 150, 151, and 153.

[487] Russell D. Johnson III. Nist standard reference database number 101. `http://cccbdb.nist.gov/`, 05-22 22/05/2022. See page 150.

[488] Simen Reine, Trygve Helgaker, and Roland Lindh. Multi-electron integrals. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(2):290–303, 2012. See page 193.

[489] Harry F King and Michel Dupuis. Numerical integration using rys polynomials. *Journal of Computational Physics*, 21(2):144–165, 1976. See page 193.

[490] Karina Viviana Rodriguez, Gustavo Gasaneo, and Dario Marcelo Mitnik. Accurate and simple wavefunctions for the helium isoelectronic sequence with correct cusp conditions. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 40(19):3923, 2007. See page 261.

[491] Eduard Matito, Jerzy Cioslowski, and Sergei F Vyboishchikov. Properties of harmonium atoms from fci calculations: Calibration and benchmarks for the ground state of the two-electron species. *Physical Chemistry Chemical Physics*, 12(25):6712–6716, 2010. See page 261.