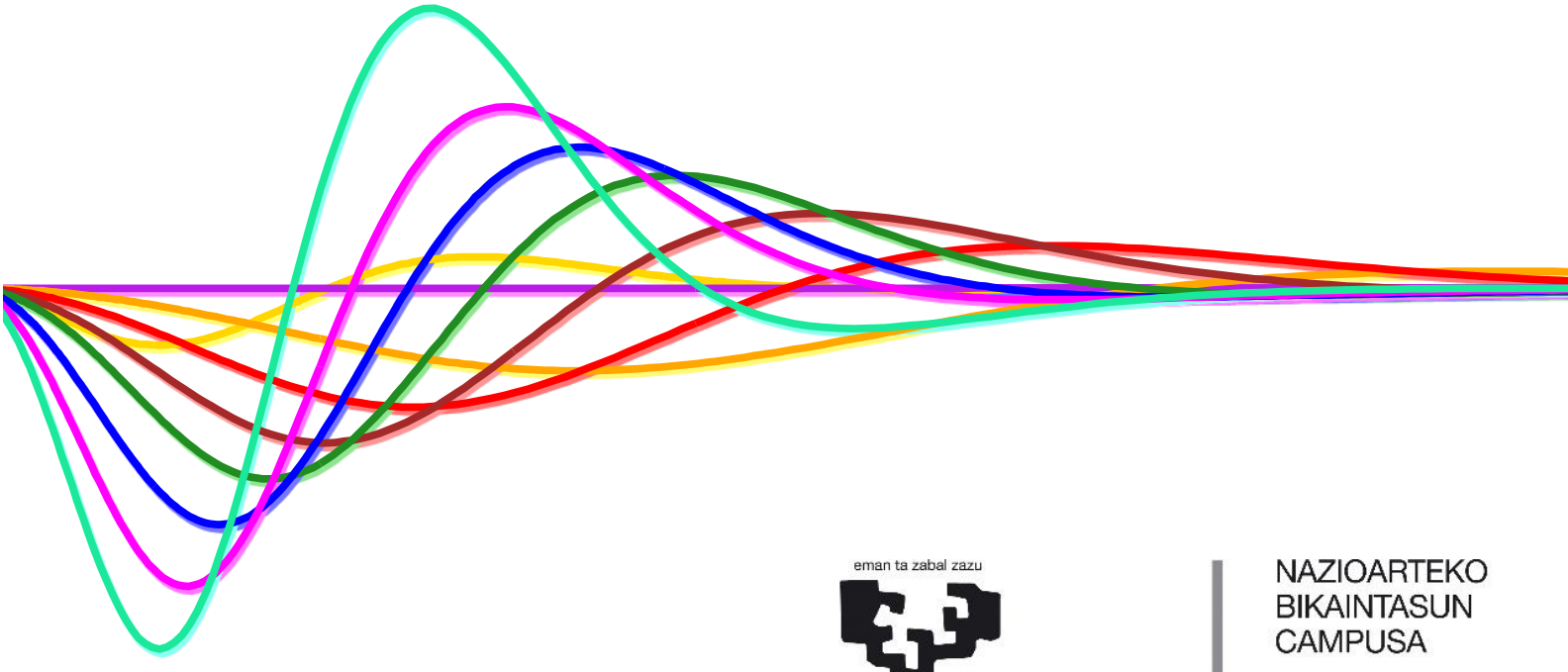
Electron correlation in density matrices and Coulomb holes

Mireia Via Nadal

2020



eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA

CAMPUS DE EXCELENCIA INTERNACIONAL

Electron correlation in density matrices and Coulomb holes

A dissertation submitted to

Universidad del País Vasco / Euskal Herriko Unibertsitatea

(UPV/EHU)

presented by

Mireia Via Nadal

2020, Donostia

Doctoral Programme in Theoretical Chemistry and Computational Modeling

Supervised by Dr. Eduard Matito Gras

Tutorized by Dr. Xabier Lopez Pestaña

(cc) 2020 Mireia Via Nadal (cc by-nc 4.0)

Dedication

Pel Sergi, que m'ha acompanyat en aquesta aventura.

Saps? M'agrada el teu país,
els prats tan verds,
d'un verd que mai no és igual
i l'herba humida
i el mar que lluita sempre
i la grisor del cel.

– A un amic del País Basc, Guillermina Motta, 1968.

Acknowledgements

Estos agradecimientos los escribo durante la cuarentena del COVID-19. No había planeado que mi último mes en la ciudad fuera a ser tan distanciado de la gente de la que quería despedirme, pero también me ayuda a valorar todas las cosas rutinarias y mundanas que he hecho con toda la gente! Quisiera agradecer todos los momentos con los del grupo de Kimika Teorikoa de la EHU: las kroketas, los kafes de las once, los de después de comer (los de la mañana en psiko no, eso es muy temprano para mi...!), las sociedades, las sagardotegis, los laboratorios... Milesker Mario, Rafa, Eli, Txema, Xabi, Elena, Vero, Txoni, Jesus, DavidC, DavidS.!

También querría agradecerselo a los estudiantes de doctorado, que han hecho mi día a día más entretenido (y más que eso): Juranga, Elisa (a ti ya te conocí doctora!), Mau, Maria, Mikel, Raul, Pep, Carmelo, Sebastian, Xiang, Diego, ... y mucha gente más entre el DIPC, el CFM y la facultad. También quería agradecerselo a los compis del piso de Agirre Miramon, Adri y Montull, y la constante música del piso! Also a special mention for Manu Fromager, and Sajan and his friends in Strasbourg, who were super hospitable and kind with me! Merci! Y creo que también se merece una línea en los agradecimientos el Mala Gissona y sus barmans! hahah os echaré de menos.

També ho vull agraïr als meus amics a Catalunya que, tot i que no els hagi pogut veure tant com m'agradaria, he sentit la seva companyia i afecte: Alba, Efri, Toni, Alex, Piro, Arantxa, Ceci, Gian, Elena. També a tu Eli, que tot i ser de Vilafranca estàs vivint per Donosti! A la gent de Química UB i les respectives reunions anuals a Llívia. També hi ha part de vosaltres dins d'aquesta tesi!

També necessito donar les gràcies a la família, sobretot a me mare i al Kiku, al Marc, a la Tinta, als meus avis i tiets i cosins. Us he trobat a faltar aquests quatre anys. Al Sergi (que ja t'he dedicat la tesi, què més vols!!), per aguantar tots els altibaixos que he tingut. A mon pare, que sempre estava orgullós de mi fes el que fes, i no s'hagués imaginat mai que hagués acabat fent un doctorat en química

teòrica.

Menció especial a l'Edu i a l'Eloy! Us vull agraïr tota la paciència i dedicació que heu tingut per instruir-me i ajudar a formar-me aquests anys! M'ho he passat molt bé treballant amb vosaltres i he après molt, en molts aspectes. Moltes, moltes gràcies! No us ho podré agraïr mai prou.

Y para terminar, querría agradecerle a un grupo de personas que han estado siempre presentes a lo largo de estos cuatro años: Irene, Maru, Ioni, Sofi, Cris, y Jorge. Sin vosotros, este capítulo de mi vida no se puede entender!

El buen ambiente que hay en el grupo de Kimika Teorikoa me hizo sentir como una más desde el minuto uno, y ha hecho que hacer la tesis sea mucho más ameno! El resumen de todo esto es que no me arrepiento para nada de haber venido a Donosti, puesto que he conocido a gente estupenda, he aprendido mucho a nivel profesional y personal, y me llevo muchas amistades!

Eskerrik asko denoi! Gero arte!

Full list of publications

This thesis is presented as a compendium of publications:

- M. Via-Nadal, M. Rodríguez-Mayorga, and E. Matito. **A Salient Signature of van der Waals interactions**, *Phys. Rev. A* **2017**, 96, p. 050501(R)
- M. Via-Nadal, M. Rodríguez-Mayorga, E. Ramos-Cordoba, and E. Matito. **Singling out Dynamic and Nondynamic Correlation**, *J. Phys. Chem. Lett.* **2019**, 10, pp. 4032–4037
- M. Via-Nadal, M. Rodríguez-Mayorga, E. Ramos-Cordoba, and E. Matito. **Range Separation of the Coulomb Hole**, *Submitted to J. Chem. Theory Comput* **2020**
- M. Via-Nadal, E. Ramos-Cordoba, P.-F. Loos, and E. Matito. **A Density Matrix Functional Approximations Benchmark for Atoms and Molecules**, *In preparation*

Publications not included in this thesis:

- M. Rodríguez-Mayorga, E. Ramos-Cordoba, M. Via-Nadal, M. Piris, and E. Matito. **Comprehensive Benchmarking of Density Matrix Functional Approximations**, *Phys. Chem. Chem. Phys.* **2017**, 19, pp. 24029–24041
- M. Rodríguez Mayorga, M. Via-Nadal, M. Solà, J. M. Ugalde, X. Lopez, and E. Matito. **Electron-Pair Distribution in Chemical Bond Formation**, *J. Phys. Chem. A* **2018**, 122, pp. 1916–1923

List of abbreviations

Abbreviation	Description
1-rDM	First–order reduced density matrix
1-DM	First–order density matrix (tensor in orbital representation)
2-PD	Pair density, or two electron density
2-rDM	Second-order reduced density matrix
2-DM	Second-order density matrix (tensor in orbital representation)
$\alpha\alpha$	Electron pairs with same spin; like-spin electron pairs
$\alpha\beta$	Electron pairs with opposite spin; opposite-spin electron pairs
ACSE	Antihermitian contracted Schrödinger equation
BBC2	Gritsenko, Pernal and Baerends functional
CA	Csányi and Arias functional
CASSCF	Complete active-space self-consistent field
CASPT2	Complete active-space second-order perturbation theory
CGA	Csányi, Goedecker and Arias functional
CI	Configuration Interaction
CC	Coupled-cluster
CCSD	Coupled-cluster including only single and double excitations
DC	Dynamic correlation
DFA	Density functional approximation
DFT	Density functional theory

E_c	Electron correlation energy
EPD	(Radial) extracule probability density
FCI	Full-Configuration Interaction
F_L	Fermi level
GU	Goedecker and Umrigar functional
HF	Hartree–Fock
HOMO	Highest occupied molecular orbital
IPD	(Radial) intracule probability density
KS	Kohn–Sham
LR	Long-range
LUMO	Lowest unoccupied molecular orbital
MBB	Müller and Baerends-Buijse functional
ML	Marques and Lathiotakis functional
MLSIC	Marques and Lathiotakis functional including self-interaction correction
MP2	Second order Møller–Plesset perturbation theory
MRCI-SD	Multireference configuration interaction with single and double excitations
NDC	Nondynamic correlation
NOrb	Natural orbital
NOcc	Natural occupation number
PNOF	Piris natural orbital functional
rDM	Reduced density matrix
rDMFT	Reduced density matrix functional theory
rDMFA	Reduced density matrix functional approximation
SD	Single determinant
SR	Short-range

T_e	Electron kinetic energy
V_{ee}	Electron-electron repulsive potential
V_{eN}	Electron-nucleus attractive potential
vdW	van der Waals

List of Figures

List of Tables

Abstract

Most of the available computational methods nowadays were designed, in general, to reproduce particular properties for particular electronic systems, being mostly accurate approximations for a given range of applicability. Electronic structure methods are nowadays very precise and accurate, yet a method that is general enough, being able to describe the properties of any system with precision and low computational cost, still does not exist. The main bottleneck in electronic structure method development is the description of the correlated motion of electrons, which hinders the computational calculation.

This thesis has focused on the study of electron correlation by considering the most used definitions to split it for method development. These components are known as dynamic and nondynamic correlation, and in this thesis they have been studied in terms of the distance between electron pairs. The means to perform such study is the intracular probability density, a distribution function that gathers the correlation information contained in the electron pair density. The intracule function is most known for its role in the definition of Coulomb holes, which provide a visual representation of the effect of including a correlated description of the electron motion in an uncorrelated framework (as the Hartree–Fock method).

An electron correlation decomposition scheme is presented, where the correlated pair density (the part of the pair density that contains all the actual correlation information) is partitioned in two components. Such components can be directly related to the dynamic and nondynamic components of electron correlation. This correlation decomposition scheme has been applied to study a set of molecules and physical models, in order to analyze the nature of the correlation components, and has permitted the identification of some characteristic behavior of each component. One of these characteristics is the particular long-range behavior of the dynamic correlation component, which presents a universal decay with the interatomic distance R^{-3}, comparable to the energetic decay of London dispersion interactions.

Moreover, special attention has been put into the reduced density matrix functional theory. Even though the theoretical background is already 50 years old, the applicability of this method is still emerging. Some approximations have been designed in this framework. A set considering some of these approximations is tested in order to determine whether they provide a physically sound description of various physical properties and how accurately they reproduce them.

Resumen

La mayoría de los métodos computacionales actuales están diseñados para reproducir propiedades particulares en sistemas electrónicos, por lo que resultan ser aproximaciones precisas, adecuadas solamente para un rango de aplicabilidad concreto. Aun así, los métodos de estructura electrónica deberían ser diseñados para reproducir la energía y las propiedades químicas de los sistemas electrónicos con un rango de aplicación más general. El principal problema, pero, es la correlación del movimiento de los electrones, que sigue siendo el principal reto de la química computacional en la actualidad.

Para mejorar el desarrollo de métodos teóricos, se ha profundizado en el estudio de la correlación electrónica, considerando los dos componentes más usados en el desarrollo de métodos. Estas componentes son la correlación dinámica y no dinámica de los electrones, las cuales se han estudiado en función de la distancia entre pares de electrones. La densidad de probabilidad intracular es una función que permite dicho estudio, siendo una función distribución que recoge la información almacenada en la función de pares de la densidad. Dicha función es usada para definir los agujeros de Coulomb de los sistemas, ya que permiten reflejar el efecto de incluir la descripción de la correlación de electrones con respecto a un caso no correlacionado (como en el método de Hartree–Fock).

En esta tesis se presenta un método para separar la correlación electrónica en dos componentes mediante un esquema de partición de correlación basado en la función de pares de la densidad electrónica. Las dos componentes que se obtienen a partir de la separación se pueden vincular directamente con las componentes dinámica y no dinámica de la correlación electrónica. Para validar dicho esquema de partición, se han realizado estudios de modelos físicos y sistemas moleculares, que a su vez han permitido obtener comportamientos característicos de cada tipo de correlación que se han repetido en todos los casos estudiados.

Una de estas características observadas es el comportamiento de la componente

de correlación dinámica con las distancias interelectrónicas grandes. Mediante la función intracular, se ha encontrado una huella universal de las fuerzas de dispersión o de London, donde la componente dinámica decae con la distancia interatómica R con un exponente de -3, R^{-3}, la cual está directamente relacionada con el comportamiento de la energía de dispersión, R^{-6}.

Se ha puesto especial atención en la teoría de la matriz (reducida) de la densidad. Aunque los fundamentos de la teoría ya tengan prácticamente 50 años, se trata de un método emergente, en el marco del cual se han propuesto un número moderado de aproximaciones de la función de pares. Se propone una serie de pruebas para determinar la validez de dichas aproximaciones, algunas directamente relacionadas con las propiedades físicas que una matriz de densidad debería cumplir.

Contents

Chapter 1

Introduction

The main goal of quantum chemistry is the resolution of the electronic Schrödinger equation to drag information from a molecular system. Unfortunately, it is impossible to solve the Schrödinger equation exactly, thus many efforts are put into finding approximations. Quantum chemistry is the science that deals with the formulation, evaluation and application of approximate solutions to the Schrödinger equation for molecular systems.

Electronic structure theory pursues to find a compromise between accuracy and computational cost to solve the nonrelativistic, electronic Schrödinger equation, since calculations bear an unaffordable increasing cost with the number of electrons. Whereas the modeling of molecules from first principles with a small number of electrons is affordable nowadays within the Born-Oppenheimer approximation, the cost becomes exorbitant when the treatment of medium-sized molecules and macromolecules is considered. Quantum chemical calculations are expensive due to the correlated motion of electrons, mostly known as electron correlation, which usually becomes more complicated to treat as the number of electrons in the system increases. Practical approximate methods usually tackle this problem by reformulating the N-electron problem as N effective independent-particle problems, such as the Hartree–Fock (HF) equations in wavefunction-based methods, or using the Kohn–Sham (KS) reference system in density functional theory (DFT). Such crucial reformulation enables first–principles modeling of molecules and materials. The HF method is usually the first approach to solve the many-body problem, yet the correlation of electron motions is greatly underestimated. Whereas it represents a small fraction to the total energy, the accurate description of the interelectronic interactions and their movement is decisive for chemistry and physics, affecting reaction barriers, binding properties, or the interactions between light and matter, amongst others.

The quest for a means to provide correct energies and properties of molecular systems at reduced computational cost has brought plenty of methods going beyond HF (as post–HF wavefunction theories, DFT, ...). However, a fully satisfactory approach has not been found yet. Because of this, method developers still seek for the best approach to describe electron correlation with the best compromise between accuracy and computational time. In the current Chapter, a definition to electron correlation is provided, along with different ways to classify it. The mathematical tools applied in this thesis to study electron correlation are presented in Chapters 1.2 and 1.3. A brief introduction to reduced density matrix functional theory is presented in section 1.2.3. Finally, an introduction to the theoretical schemes to decompose electron correlation is presented in Chapter 1.4.

1.1 Electron correlation

In quantum chemistry, the electronic, nonrelativistic Hamiltonian contains a pairwise Coulomb repulsion that causes the motion of electrons to be dependent on each other. It is said, then, that their motion is correlated [1, 2]. The electron correlation energy (E_c) is, however, a theoretical concept: it is not an observable linked to any operator, but rather a quantity born from the inability of the HF method to correctly describe the correlated motion of electrons. In 1933, Wigner and Seitz mentioned the concept of "electron correlation" for the first time in the quantum chemistry literature [3,4], and thereafter many scientists referred to it yet without a clear definition being given; even some scientists, such as Slater, complained about the confusion in the community due to the multiple definitions attributed to this term [5]. It was not until 1959 that Löwdin provided a solid definition to E_c in which, for a given state, E_c is the energy difference between the exact eigenvalue of the nonrelativistic electronic Hamiltonian and its expectation value in the HF approximation in a complete basis [6],

$$E_c = E_{\text{exact}} - E_{\text{HF}}; \tag{1.1}$$

that is, it is a fraction of the total, nonrelativistic energy that a single determinant wavefunction is not able to describe. In practice, the exact energy is not known, and the E_{exact} term refers to the full configuration-interaction (FCI) energy computed for a given one-electron basis set. Besides Löwdin's proposal, more definitions for E_c appeared in the literature. Some authors suggested using the unrestricted formalism of HF (UHF) instead of the restricted one (RHF), as the latter completely fails in the description of stretched geometries [7]. Instead, UHF provides an improved

behavior, but the approach still lacks electron correlation into its description. Other authors use the concept of electron correlation to quantify the amount of correlation captured by a given method compared to HF [1, 8, 9]. In KS DFT, the correlation energy can be defined as the difference between the exact energy and that of the non-interacting system [10]. Notwithstanding, Löwdin's definition is the most used and extended in the community.

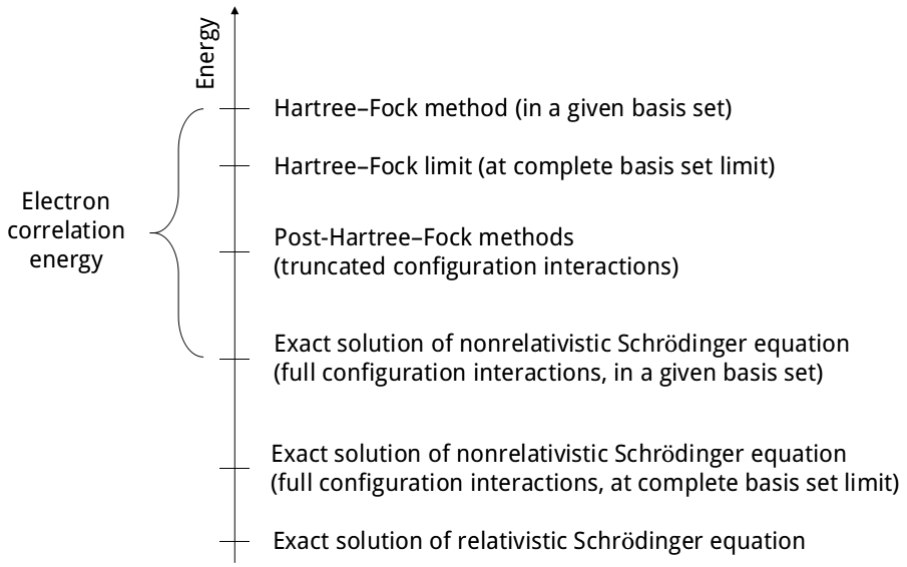


Figure 1.1: Schematic representation of Löwdin's definition of electron correlation energy. Adapted from Wikipedia, https://commons.wikimedia.org/wiki/File:Electron_correlation.png.

According to the variational principle, any post-HF variational method will yield an energy approximation closer to the exact solution for the nonrelativistic Schrödinger equation than the HF method (Fig. 1.1). Electron correlation is, then, a phenomenon lying beyond the HF approximation in which, according to Eq. 1.1, the HF description neglects the correlated motion of electrons. This statement is not completely correct, as the HF approximation does include a certain amount of correlation for same-spin electron pairs (exchange interactions) as its wavefunction preserves the fermionic nature of electrons. See Chapter 1.2.2 for more details on the HF method.

Choosing a basis set is also essential in the correct description of electron correlation. The lack of basis functions, the use of small angular momentum functions, or an incorrect preexponent value may prevent electrons to be found in positions that stabilize the molecule, leading to too repulsive configurations. The effect of the basis set chosen is already considered in Löwdin's definition because the energies are considered to be obtained at the complete basis set limit (Fig. 1.1).

Despite its deficiencies, the HF method is usually able to capture up to 99% of the total energy of a system. To put a couple of examples, it yields 96% of H_2 energy, and even the 98.9% of the helium dimer [1]. Such high percentual ratios are, however, not enough for the correct description of the electronic properties of a system: the remaining E_c is often of the order of the energy of most elemental chemical reactions. In the HF description, the dissociation of the hydrogen molecule is incorrect, with the dissociated hydrogen atoms having ionic character. Electronic structure methods, then, must accomplish the description of E_c because, otherwise, binding properties, geometries, excitation energies, reaction barriers, etc. cannot be correctly predicted. Through the years, many efforts have been put into understanding the nature of electron correlation and further development of approximate methods [11–23]. Different flavors of electron correlation are introduced in the forthcoming section.

1.1.1 Types of electron correlation

Electron correlation is a complex phenomenon, and, because of this, scientists have provided a plethora of terminology to obtain a clearer view for the analysis of the effect of the motion of electrons. Depending on the context, different jargon is used for atomic and molecular systems. Fig. 1.2 includes a graphical summary of some of the words used to separate electron correlation. In atoms, quantum chemists distinguish between radial and angular correlation because of their spherical symmetry [24, 25]. Radial correlation prevents electrons to be near an electron that is close to the nucleus, and consequently forces them to be found far from it. This correlation is also named in-out correlation and can be dealt with using basis functions with the same angular momentum but different exponent: for an electron found at the $1s$ orbital, $2s$ and $3s$ functions should be included to treat radial correlation. On the other hand, angular correlation refers to the increased probability of an electron to be found at a side of the nucleus when another one is at the opposite side. Functions with high angular momentum can describe such correlation.

In diatomic molecules, correlation is usually discerned between axial and angular. The former is also referred as left-right or longitudinal, and describes the increased probability of an electron to be located in the region of nucleus A when another electron is around nucleus B. It is also sometimes defined as the tendency of an electron to be far from the internuclear axis (the axis where the bond is located) when another electron is close to it [11, 15]. This type of correlation becomes im-

portant at large nuclear separations, such as in dissociations and transition states. To include this correlation, electrons must be allowed to be located at antibonding orbitals and describe the system with a multiconfigurational wavefunction. Angular correlation is also referred as equatorial, and is a mix of the in-out and radial correlation in atoms, being notorious when the distance between nuclei is reduced [11,26].

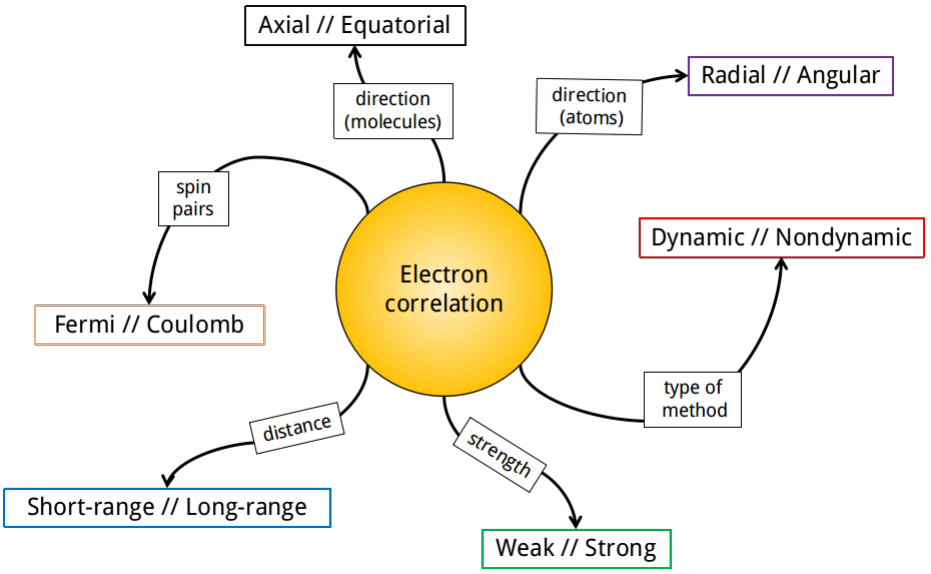


Figure 1.2: Summary of all the electron correlation jargon presented in this section, grouped according to the criterion used to separate electron correlation.

In polyatomic molecules it becomes unhandy to express correlation in terms of the location of electrons within a molecule. Instead, more general criteria are applied to separate and define electron correlation contributions. The effect of correlation according to the interelectronic distance is usually employed and discerns between long- (LR) and short-range (SR) correlation. LR interactions arise in orbital degeneracies in the dissociation limit, or are caused by the van der Waals forces, whereas SR correlation can be caused by near-degeneracies or simply by the actual movement of electrons.

Another classification comes after the definitions of Coulomb and Fermi holes, which will be introduced in Chapter 1.3.1. Electrons being charged entities, electrostatic or Coulombic repulsive interactions are forces that also correlate the electronic motion, and are an important source of electron correlation, which is called Coulomb correlation. On the other hand, Fermi (or exchange) correlation arises because electrons are fermions with two possible spin states, either $s_\alpha = \frac{1}{2}$ or $s_\beta = -\frac{1}{2}$ [1,27]. Indeed, two electrons with the same spin experience a big repulsion caused by the Pauli exclusion principle, and consequently their motion is correlated. By nature,

Fermi interactions are stronger than electrostatic ones. The jargon of Fermi and Coulomb correlation is typically (but not exclusively) used in electron correlation holes (see Chapter 1.3.1).

The distinction between dynamic and nondynamic correlation is more widely established in the community [1, 28, 29]. Electrons are said to be dynamically correlated because they are charged entities in constant motion that repel each other. Hence, any system with more than one electron (generally) presents dynamic correlation, becoming more important when the number of electrons is large. Because of this, this type of correlation is universal [30]. Considering a configuration interaction (CI) wavefunction, dynamic correlation emerges when the single determinant picture (the HF determinant) is consistent but incomplete, and the CI wavefunction is formed by a highly-contributing HF configuration and by a large amount of configurations that do not mix strongly with it (see Fig. 1.3) [28–30]. Because the HF determinant is the dominant configuration in the CI expansion, the electron density of a system with dynamic correlation will resemble the HF density. Besides, the low-contributing determinants mixing with the HF determinant will cause small and local changes to the HF density.

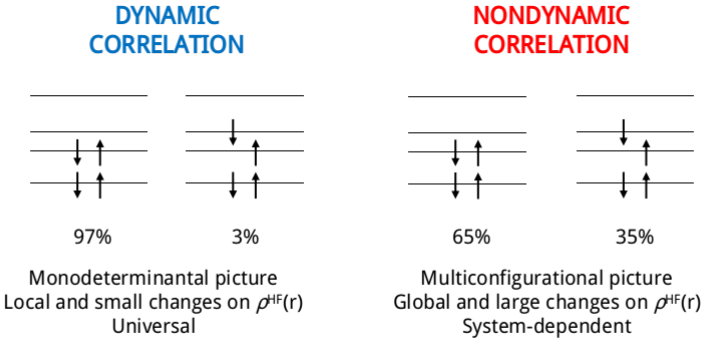


Figure 1.3: Summarized definitions of the dynamic and nondynamic types of electron correlation.

In contrast, nondynamic correlation rises when a system has (near-)degeneracies and thus the ground state cannot be described by a single determinant but with the mixing of low-lying-in-energy excited states along with the HF configuration. Then, the CI vector is formed by a short expansion of Slater determinants, and expansion coefficient of the HF determinant is relatively smaller than one. The large mix of configurations causes global changes to the HF uncorrelated density [28, 30–32], as the electron density is proportional to the square of the expansion coefficients of the CI vector and, hence, the electron density of the system differs from the HF one. In contrast to dynamic correlation, nondynamic correlation is system-dependent and,

therefore, not universal. It emerges when orbital degeneracies exist, caused, for instance, by a molecular dissociation; it appears in charge transfers, in poliradicaloids or by LR correlation caused by electron entanglement.

The distinction between dynamic and nondynamic correlation explained in the former paragraph is done in terms of the expansion coefficients in the CI vector. Strictly speaking, the criterion to distinguish between both types of correlation should be considered in the Nth root of the coefficients, N being the number of electrons. The condition for a system that is dominated by dynamic correlation is that such Nth root should be close to one [33]. For instance, consider an infinite number of noninteracting He atoms. Its CI vector is composed by N highly contributing configurations, which does not correspond to the definition provided for dynamic correlation. However, each He atom presents some dynamic correlation, and they do not give rise to nondynamic correlation because they do not interact among them. By considering the Nth root of the coefficients, which all of them squared sum one, there is no controversy in the distinction between both types of correlation since the quantity is normalized with respect to the number of particles in the system.

Some authors have introduced different nuances and provided a further distinction within nondynamic correlation. The "nondynamic correlation" term is sometimes restricted to the correlation arisen in long-distance electron interactions (as a molecule correctly dissociating into two fragments), and "static correlation" is used for the correlation arisen when a suitable combination of determinants is used to account for proper spin symmetries [27,28,32,34]. Further distinctions to nondynamic correlation were suggested, such as the one proposed by Hollett and Gill based on the ability of UHF to account for it. They identified two classes of nondynamic correlation, named types A and B. Type A arises from molecular dissociations, for instance, and it can be described by UHF. This type of correlation can be detected with the HOMO-LUMO gap energies, where the bond cleavage produces absolute near-degeneracies in the gap. Instead, type B cannot be described by UHF and it is present in systems with relative near-degeneracies [35].

In the quantum chemistry field, it has been customary to associate dynamic correlation with SR interactions and nondynamic correlation with the LR ones. It was accepted that the inability to model the electron-electron cusp avoids the correct description of dynamic correlation [36, 37]; on the other hand, nondynamic correlation is linked to LR interactions because electrons tend to follow their re-

spective atom as a molecule dissociates, which gives rise to nondynamic correlation. This former conception changed and nowadays the two terminologies have been disconnected. Whereas usually nondynamic(dynamic) correlation indeed arises in large(short) interelectronic separations, one cannot exclusively associate them to those. For instance, it is considered that London dispersion interactions are a type of LR dynamic correlation [38, 39]. Information regarding dispersion interactions is expanded in section 1.1.2.

The separation of electron correlation into dynamic and nondynamic correlation contributions has become the most widely used, for it permits a qualitative classification of electronic structure methods according to their ability to describe these correlation types. The correct treatment of nondynamic correlation requires a multi-configurational method to strongly mix a small number of other configurations with the HF one, such as in the complete active-space self-consistent field method (CASSCF) [40] or the density matrix renormalization group (DMRG) [41]. Instead, the second-order Møller–Plesset perturbation theory (MP2) [42], configuration interactions with single and double excitations (CISD) [43], or coupled cluster with single and double excitations (CCSD) [44] introduce dynamic correlation via a small mixing of a large number of configurations in the wavefunction [45]. The universal character of DFT approximations also permits the inclusion of dynamic correlation to some extent [15]. Whereas there is a plethora of methods to treat both types of correlation separately, very few approaches are able to depict both kinds correlation contributions simultaneously, being CASPT2 [46, 47] (a mix of CASSCF with a posterior MP2 treatment) a textbook example. Other examples are the multi-reference single and double configuration interaction method (MRCI-SD) [48, 49], the antihermitian contracted Schrödinger equation (ACSE) [50], or, most recently, the adiabation-connection MCSCF (AC-MCSCF) [51] and the ΔNO [52].

These methods are able to capture electron correlation accurately, yet neither of them presents a convenient scaling of the computational cost with the system size. Because of this, many hybrid schemes have been proposed, in which different methods to tackle both correlation types have merged to produce accurate results but at a more reasonable cost [53–55]. An example are range-separated approaches [56], which have gained importance in the last decade. Electronic structure methods that recover separately different correlation types are combined through a range-separation function so SR interactions include dynamic correlation, and LR asymptotics is modeled accordingly with methods that include nondynamic correlation effects [57, 58].

Through this Chapter, a clear distinction has been made between dynamic and nondynamic electron correlation contributions, which are presented as independent phenomena where electronic structure methods have been classified accordingly to their ability to retrieve them. However, notice that there is not a clear-cut separation between dynamic and nondynamic correlation since it is a matter of mixing of states being more or less contributing to the CI vector [59]. In the end, the sum of both correlation types recovers the complete electron correlation, $E_{\mathrm{c}} = E_{\mathrm{DC}} + E_{\mathrm{NDC}}$. Hence, post-HF methods are able to describe not only an exclusive type of correlation but both, yet to a different extent. For instance, a CISD calculation of a given system will basically account for dynamic electron correlation, but it will also introduce a certain small amount of nondynamic correlation in the wavefunction. However, it is useful to grossly separate the two types of electron correlation in order to develop electronic structure methods.

1.1.2 Dispersion interactions

In this thesis, special attention has been put into dispersion interactions, an electron correlation phenomena arising from dynamic correlation. They are originated from London dispersion or van der Waals (vdW) forces, and consequently weakly correlate the motion of electrons [60]. They are a type of noncovalent interactions (NCIs) and thus are originated from electrostatic interactions formed from constantly fluctuating electron clouds between two or more fragments. Therefore, they are many-body and nonlocal intermolecular forces that arise when at least two fragments are largely separated. Note that in this thesis we do not apply the terminology usually used in chemistry books, where the term "van der Waals interactions" embraces the electrostatic, induction and dispersion interactions. Hereafter, the aforementioned forces will be referred to as NCIs, and the term "van der Waals interactions" will be used indistinguishably with "dispersion London interactions".

Fig. 1.4 has been included to provide a better understanding of how dispersion interactions are originated. The classical electrostatic interaction between two fragments with permanent dipoles (as in two water molecules) can be represented with the third picture alone in Fig. 1.4. Instead, if one of those fragments does not present a permanent dipole, the electric field generated by the polar fragment will distort the polarless fragment, causing an induced dipole. This event is described by the second and third plots in Fig. 1.4. For dispersion interactions, consider that none of the fragments has a permanent dipole. The constant fluctuation of electrons causes an instantaneous, transient charge distribution in one fragment that, at the

same time, induces another dipole on the neighboring fragment. This couples the two instantaneous dipoles, a process that corresponds to the complete scheme represented in Fig. 1.4 [60].

These short-lived, complementary dipoles cause a weak and nonlocal electrical attraction in which the energetically stable distance depends on the size of each fragment's electron cloud. If the clouds of adjacent fragments are too close, the attracting dispersion interaction becomes repulsive and other electrostatic forces become prevalent. Note that, compared to covalent interactions, NCIs imply an extremely small contribution to the total energy and are less directional and distance-dependent [61]. London dispersion interactions are also known as vdW interactions in honor of van der Waals, who first introduced attractive forces between molecules in his equation of state [62].

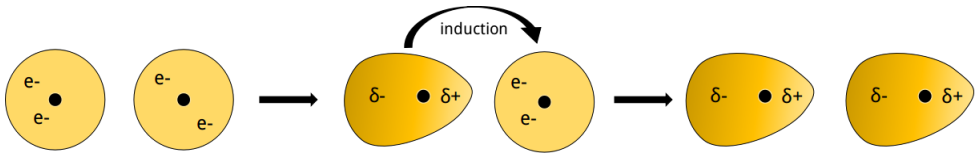


Figure 1.4: Graphic representation of noncovalent interactions. Yellow areas represent electron densities. The third figure alone represents two permanent dipoles interacting via electrostatic forces. The second and third figures together describe a dipole inducing another dipole to a polarless atom. The complete figure depicts a polarless atom in which the electron distribution generates an instantaneous dipole (first to second figure), which in turn induces another dipole to another polarless atom (second to third figure).

These serendipitous interactions may seem irrelevant and very weak, yet they are of great importance for the structure, stability, dynamics and functionality of electronic systems. Their origin causes them to be ubiquitous in nature, and play an important role in chemistry, physics, and biology determining the properties from the microscale to the macroscale. To name some examples, vdW interactions are responsible for the binding properties in molecules and materials [63–66], the crystal formation in molecules, the physics of layered materials and their cohesive interactions, as in graphite [67, 68] (yet some authors have claimed that they are bound by weak metallic forces) [69], the double-helix structure of DNA, as well as many other features in proteins such as their structures, protein-protein recognition, or drug binding [70, 71]. They are also responsible for the cohesion in asteroids [72], and they are involved in the mechanism that gives geckos the ability to adhere to walls [73–75].

Dispersion forces were initially thought to result from the polarization of a fragment in the field of a permanent dipole or quadrupole moment from a neighboring fragment [76, 77], and scientists were not able to explain the nature of dispersion interactions until quantum mechanics was developed. In a classical treatment, electrostatic interactions are usually dissected and analyzed by means of a multipole expansion [60]; however, their non-classical origin obliges the treatment of dispersion interactions by means of perturbation theory [60, 78, 79], first accomplished by London [38, 80]. The main assumption is that electron clouds of two fragments do not overlap when $R \to \infty$, being R the interfragment separation. There, the excitations on both monomers are coupled and permit rewriting the dispersion energy in terms of the dynamical polarizabilities of each monomer. One can follow the development specified in references [60, 78], where it is demonstrated that the dispersion energy takes the form of

$$E_{\text{disp}} = -\frac{3U_A U_B}{2(U_A + U_B)} \frac{\alpha_A \alpha_B}{(4\pi\epsilon_0)^2 R^6} = -\frac{C_6}{R^6}, \tag{1.2}$$

where A and B denote the fragment, U_A and U_B are average excitation energies of fragments A and B, α denotes the atomic dynamic polarizability, and C_6 is the sixth-order dispersion coefficient. Higher powers can be added to the leading R^{-6} term if higher perturbation orders are considered on the treatment, such as $C_8 R^{-8}$, $-C_{10} R^{-10}$ and $C_{12} R^{-12}$, where the latter term is used to describe the SR interaction in the Lennard-Jones potential [81]. The energetic dependency with R is not always dominated by the R^{-6} leading term. Dispersion forces present a dependence with light that affects the forces felt by matter when light is irradiated. If the interfragment distance R is larger than the characteristic absorption frequency of the molecule λ, a retardation of the electromagnetic field generated by the electron clouds arises (retardation effects). When this happens, the dispersion-type emerging forces are called Casimir forces and changes the proportionality of the dispersion energy to R^{-7} [60, 82–84].

Since there is no way to determine the excitation energies nor the accurate fragment polarizabilities within a molecule, dispersion coefficients C_i are usually obtained from approximated treatments in the Casimir–Polder formula (*vide infra*). In 1948, Casimir and Polder proposed a perturbation theory treatment that is approached in terms of quantum electrodynamics, and provides an exact definition for dispersion interactions [83]. The adiabatic-connection fluctuation-dissipation (ACFD) theorem is applied for the E_{c} and permits a description in terms of the density-density response function at an imaginary frequency, $\chi(\boldsymbol{r_1}, \boldsymbol{r_2}, i\nu)$, which is generally approximated to two local atomic polarizability densities. The Casimir–

Polder formula reads, in atomic units,

$$E_{\text{disp}} = -\frac{3}{2\pi}\int_0^\infty d\nu \iint d\boldsymbol{r_1} d\boldsymbol{r_2} \alpha(\boldsymbol{r_1}, i\nu)\alpha(\boldsymbol{r_2}, i\nu)\frac{f(R)^2}{R^6}, \tag{1.3}$$

where $f(R)$ is a damping function used to separate the energy by ranges [60, 85]. From this equation, the C_6 coefficient is approximated to

$$C_6 = \frac{3}{\pi}\int_0^\infty d\nu \iint d\boldsymbol{r_1} d\boldsymbol{r_2} \alpha(\boldsymbol{r_1}, i\nu)\alpha(\boldsymbol{r_2}, i\nu). \tag{1.4}$$

The application and features of the Casimir–Polder equation are out of the scope of this thesis. More details can be found in references [60, 66, 85, 86].

1.1.2.1 van der Waals-including methods

Because vdW interactions are an electron correlation effect, their modeling has been a challenge and a central topic in the community [87]. Wavefunction-based methods become a bad choice because a high level of theory is required, leading to prohibitively high computational costs. On the other hand, most density functional approximations (DFAs) are known to provide good SR dynamic correlation, yet they fail to provide the correct decay in the LR part of the electronic energy due to their semilocal nature [88]. Many accurate first–principles approaches have flourished in recent years, but some of them are still computationally demanding and may present some deficiencies. There are plenty of reviews present in the literature [18, 19, 66, 85, 89–91], in which the vdW-treating methods are essentially classified in four general groups:

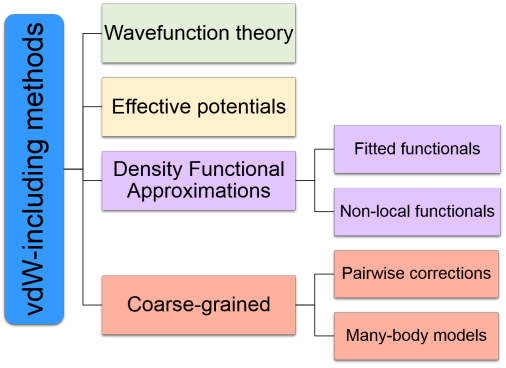


Figure 1.5: Schematic summary of the vdW-including methods available.

1. Effective, atom-centered potentials that are fitted to reproduce NCIs [92].

2. Wave-function theory methods as the quantum Monte-Carlo method [22], the coupled cluster (CC) method [93], and random-phase approximation (RPA) approaches [89, 94, 95] to solve the Casimir–Polder equation (Eq. 1.3), which account for a great accuracy in the description of dispersion interactions but with the penalty of large computational costs.

3. Density functional approximations (DFAs), subdivided into two groups:

   3.1. DFAs that have been fitted to reproduce NCIs and its consequent LR asymptotics, especially the Minnesota functionals [96].

   3.2. DFAs that include a nonlocal kernel that permits the correct LR asymptotics. Roughly speaking, these methods are first–principles based — that is, they explicitly depend on the electron density or orbitals. This group gathers the van der Waals exchange-correlation functionals (vdW-DF and vdW-DF2) [97, 98], the Vydrov and Van Voorhis nonlocal functionals VV09 and VV10 [99, 100], or the damped asymptotic dispersion energy (DADE) functional [86].

4. Coarse-grained approaches, which can also be divided into two subgroups:

   4.1. Pairwise, additive corrections to the total energy, as in the exchange-hole-dipole moment model (XDM) [101–106], the Tkatchenko-Scheffler model (DFT+TS) [107], and Grimme's dispersion corrections, the widely used DFT-D3 and the recently developed DFT-D4 [108, 109].

   4.2. Many-body coarse-grained models, such as the many-body dispersion (MBD) scheme [110, 111], and the DFT/vdW-QHO-WF, which adopts the Quantum Harmonic Oscillator (QHO) model and is based in maximally localized Wannier Functions (WF) [112].

Nowadays, there exist plenty of approaches that permit the treatment of van der Waals interactions in electronic structure calculations. The D3 additive correction scheme has become an everyday tool in computational chemistry thanks to its easy applicability and computational efficiency, bringing a practically negligible cost to large systems. The approach considers the molecular geometry as the only input and includes a correction to the total energy that is added to the result of a conventional DFA calculation. The recently presented D4 model has also introduced the atomic partial charges as an input. The energy correction influences geometries and frequencies, as they are connected to the derivatives of the energy with respect to the atomic coordinates. However, it does not influence all the electronic properties in a molecule. Particularly, consider the D3 dispersion correction for an excited state

with a given geometry. The correction for this state would be very similar to the corresponding correction for the ground state with the same geometry (not entirely due to the effect of the atomic partial charges, which are also considered). In a sense, additive corrections to the energy are added *a posteriori* and, therefore, dispersion interactions are not treated self-consistently. Because of this, further research and development are required on first–principles-based methods in which the dispersion interactions can be optimized in the self-consistent procedure [113, 114].

1.2 Theoretical framework

Density matrices are used to study electron correlation in quantum physics and chemistry. In 1960, Coulson already stated that "electron correlation should show up in the two-particle density matrix" [115]. Indeed, it is well known that the energy is an exact functional of the pair density, being a central quantity to describe electron correlation. Hence, a brief introduction is given through this Chapter.

1.2.1 Density matrices

Within the Born-Oppenheimer approximation, the nonrelativistic, quantum-mechanical electronic Hamiltonian operator for a system composed of N electrons and M nuclei is composed of three operators. In atomic units,

$$\widehat{H} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{a=1}^{M}\sum_{i=1}^{N} \frac{Z_a}{|\boldsymbol{R}_a - \boldsymbol{r}_i|} + \sum_{i<j}^{N} \frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}, \tag{1.5}$$

where each term is, in order of appearance, the electronic kinetic energy operator T_e, the electron-nucleus interaction operator V_{eN}, and the interelectronic repulsion operator V_{ee}, with Z_a being the charge of a given nucleus a. The total energy is obtained when the nucleus-nucleus attraction operator V_{NN} is included. Due to the character of the operators, the electronic Hamiltonian can be separated in one- ($h(i)$) and two-electron ($v(i,j)$) operators,

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_a \frac{Z_a}{|\boldsymbol{R}_a - \boldsymbol{r}_i|}; \tag{1.6}$$

$$v(i,j) = \frac{1}{|\boldsymbol{r_i} - \boldsymbol{r_j}|}, \tag{1.7}$$

so the electronic Hamiltonian can be expressed simply as

$$\hat{H} = \sum_i h(i) + \sum_{i<j} v(i,j). \tag{1.8}$$

In electronic structure methods, the wavefunction for a given system is obtained as an approximate solution for the time-independent Schrödinger equation,

$$\widehat{H}\Psi(\mathbf{1},\mathbf{2},\ldots,\boldsymbol{N}) = E_i\Psi(\mathbf{1},\ldots,\boldsymbol{N}) \tag{1.9}$$

where E_i is the electronic energy for state i, and $\Psi(\mathbf{1},\ldots,\boldsymbol{N})$ is the wavefunction of the system, and $\mathbf{1}$ are the spin and three-dimensional coordinates of an electron, namely $\mathbf{1} \equiv (\boldsymbol{r_1}, \sigma_1)$, and therefore the same nomenclature is used for the differentials , $d\mathbf{1} = d\boldsymbol{r_1}d\sigma_1$.

The wavefunction is a $4N$-variable object, involving spin and spatial coordinates, so its interpretation may become arduous. Instead, for the purpose of quantum chemistry, density matrices provide a more interpretable point of view. Firstly introduced by von Neumann in 1927 to describe statistical concepts in quantum mechanics [116], it was not until 1930 that Dirac applied density matrices within the quantum mechanical context [117]. Compared to wavefunctions, density matrices permit to express the maximum information available of a system in a compact manner, avoiding the introduction of unnecessary variables. They allow the description of a quantum system in a mixed state, being a more complete object than the wavefunction is [118]. One can define the Nth order density matrix (or simply N-density matrix) of an N-electron wavefunction as

$$\rho_N(\mathbf{1'},\mathbf{2'},\ldots,\boldsymbol{N'};\mathbf{1},\mathbf{2},\ldots,\boldsymbol{N}) = \Psi^*(\mathbf{1'},\mathbf{2'},\ldots,\boldsymbol{N'})\Psi(\mathbf{1},\mathbf{2},\ldots,\boldsymbol{N}), \tag{1.10}$$

being a tensor formed by $2N$ variables [119,120]. From the Nth density matrix, one can obtain lower-rank matrices by integrating the appropriate electronic coordinates, which are called n-order reduced density matrices (n-rDM),

$$\rho_n(\mathbf{1'},\ldots,\boldsymbol{n'};\mathbf{1},\ldots,\boldsymbol{n}) = \binom{N}{n} n! \int \rho_N(\mathbf{1'},\ldots,\boldsymbol{N'};\mathbf{1},\ldots,\boldsymbol{N}) d(\boldsymbol{n}+\mathbf{1})\ldots d\boldsymbol{N} d(\boldsymbol{n}+\mathbf{1})'\ldots d\boldsymbol{N'} \tag{1.11}$$

being $n < N$, and the so-called McWeeny normalization $\binom{N}{n}n!$ is considered, $\mathrm{Tr}[\rho_n] = \binom{N}{n}n!$ [118]. Other normalization factors are applied throughout the literature, such as the Löwdin normalization, with $\binom{N}{n}$ [121]. Lower-rank densities (m) can be obtained from higher-rank densities (l) by integration,

$$\rho_m(\mathbf{1'},\ldots,\boldsymbol{m'};\mathbf{1},\ldots,\boldsymbol{m}) = \frac{1}{(N-m)\ldots(N-l+1)} \int \rho_l(\mathbf{1'},\ldots,\boldsymbol{m'},(\boldsymbol{m}+\mathbf{1})',\ldots,\boldsymbol{l'};\mathbf{1},\ldots,\boldsymbol{l}) d(\boldsymbol{m}+\mathbf{1})\ldots d\boldsymbol{l}\, d(\boldsymbol{m}+\mathbf{1'})\ldots d\boldsymbol{l'}, \tag{1.12}$$

but not *vice versa*: the integration step reduces the information enclosed in a density matrix, and the information originally contained cannot be recovered unless the wavefunction is known. Density matrices are Hermitian

$$\rho_n(\mathbf{1'},\ldots,\boldsymbol{N'};\mathbf{1},\ldots,\boldsymbol{N}) = \rho_n^*(\mathbf{1},\ldots,\boldsymbol{N};\mathbf{1'},\ldots,\boldsymbol{N'}) \tag{1.13}$$

and antisymmetric in the exchange of a pair of indices

$$\begin{aligned}\rho_n(\ldots,\boldsymbol{i'},\ldots,\boldsymbol{j'},\ldots;\ldots,\boldsymbol{i},\ldots,\boldsymbol{j},\ldots) = \\ -\rho_n(\ldots,\boldsymbol{j'},\ldots,\boldsymbol{i'},\ldots;\ldots,\boldsymbol{i},\ldots,\boldsymbol{j},\ldots).\end{aligned} \tag{1.14}$$

The first– and second–order rDMs (1-rDM and 2-rDM, respectively) are widely used in quantum mechanics. They are defined as

$$\begin{aligned}\rho_1(\mathbf{1'};\mathbf{1}) =& N\int \rho_N(\mathbf{1'},\ldots,\boldsymbol{N'};\mathbf{1},\ldots,\boldsymbol{N})\delta(\mathbf{2}-\mathbf{2'})\ldots\delta(\boldsymbol{N}-\boldsymbol{N'})d\mathbf{2}\ldots d\boldsymbol{N}d\mathbf{2'}\ldots d\boldsymbol{N'} \\ =& \frac{1}{N-1}\int \rho_2(\mathbf{1'},\mathbf{2'};\mathbf{1},\mathbf{2})\delta(\mathbf{2}-\mathbf{2'})d\mathbf{2}d\mathbf{2'}\end{aligned} \tag{1.15}$$

and

$$\begin{aligned}&\rho_2(\mathbf{1'},\mathbf{2'};\mathbf{1},\mathbf{2}) \\ &\quad =N(N-1)\int \rho_N(\mathbf{1'},\ldots,\boldsymbol{N'};\mathbf{1},\ldots,\boldsymbol{N})\delta(\mathbf{3}-\mathbf{3'})\ldots\delta(\boldsymbol{N}-\boldsymbol{N'})d\mathbf{3}\ldots d\boldsymbol{N}d\mathbf{3'}\ldots d\boldsymbol{N'} \\ &\quad =\int \Psi^*(\mathbf{1'},\ldots,\boldsymbol{N'})\Psi(\mathbf{1},\ldots,\boldsymbol{N})\delta(\mathbf{3}-\mathbf{3'})\ldots\delta(\boldsymbol{N}-\boldsymbol{N'})d\mathbf{3}\ldots d\boldsymbol{N}d\mathbf{3'}\ldots d\boldsymbol{N'},\end{aligned} \tag{1.16}$$

respectively. The normalization factor for the 2-rDM corresponds to the number of electron pairs, whereas the 1-rDM is normalized to the number of electrons in the system.

The diagonal elements of density matrices are the so-called n-electron densities,

$$\rho_n(\mathbf{1},\ldots,\boldsymbol{n}) = \int \rho_n(\mathbf{1'},\ldots,\boldsymbol{n'};\mathbf{1},\ldots,\boldsymbol{n})\delta(\mathbf{1'}-\mathbf{1})\ldots\delta(\boldsymbol{n'}-\boldsymbol{n})d\mathbf{1'}\ldots d\boldsymbol{n'}, \tag{1.17}$$

with $n \leq N$. n-electron densities measure the joint probability of finding n electrons regardless of the position in space of the rest of $N-n$ electrons. By virtue of their corresponding density matrix, the electron densities are symmetric in the exchange of a pair of indices, and inherit the normalization factors from their corresponding density matrix counterpart. The most well-known n-density in the chemistry community is the one coming from the 1-rDM, the 1-electron density, $\rho(\mathbf{1}) \equiv \rho_1(\mathbf{1};\mathbf{1})$, which is simply called electron density, and is the main quantity in DFT. The diagonal elements of the 2-rDM are known as the 2-electron density, pair probability

density, pair density, or 2-particle density (2-PD), and define the probability of finding a couple of electrons irrespectively to where the rest $N-2$ are.

Quantum chemical calculations work with basis sets, thus it is appropriate to use density matrices in the orbital representation. Any (reduced) density matrix can be expressed as an expansion of a set of orbitals of size K, $\{\phi_i(\mathbf{1})\}_{i=1,K}$,

$$\rho_n(\mathbf{1'},\ldots,\boldsymbol{n'};\mathbf{1},\ldots,\boldsymbol{n}) = \sum_{\substack{i_1\ldots i_n \\ j_1\ldots j_n}}^{K} {}^n\mathrm{D}^{i_1\ldots i_n}_{j_1\ldots jn}\phi^*_{i_1}(\mathbf{1'})\ldots\phi^*_{i_n}(\boldsymbol{n'})\phi_{j_1}(\mathbf{1})\ldots\phi_{j_n}(\boldsymbol{n}), \quad (1.18)$$

and the analogous expression can be written for the spinless n-rDM and the corresponding n-densities. The orbital representation of an rDM is described by the ${}^n\mathrm{D}$ tensor, named n-density matrix (n-DM). Whereas the spinless n-rDM is composed by $3n$ variables, the n-DM is a $2n$-index tensor of dimension K^n. The orbital representations for the aforementioned spinless 1-rDM and 2-rDM are

$$\rho_1(\boldsymbol{r'_1};\boldsymbol{r_1}) = \sum_{ij}^{K} {}^1\mathrm{D}^i_j\phi^*_i(\boldsymbol{r'_1})\phi_j(\boldsymbol{r_1}) = \sum_{i}^{K} n_i\chi^*_i(\boldsymbol{r'_1})\chi_i(\boldsymbol{r_1}) \quad (1.19)$$

$$\begin{aligned}\rho_2(\boldsymbol{r'_1},\boldsymbol{r'_2};\boldsymbol{r_1},\boldsymbol{r_2}) &= \sum_{\substack{ij \\ kl}}^{K} {}^2\mathrm{D}^{ij}_{kl}\phi^*_i(\boldsymbol{r'_1})\phi^*_j(\boldsymbol{r'_2})\phi_k(\boldsymbol{r_1})\phi_l(\boldsymbol{r_2}) \\ &= \sum_{i,j}^{K} \gamma_{ij}\Gamma^*_i(\boldsymbol{r'_1},\boldsymbol{r'_2})\Gamma_j(\boldsymbol{r_1},\boldsymbol{r_2}) \end{aligned} \quad (1.20)$$

The *r.h.s.* of both Eqs. 1.19 and 1.20 define the corresponding density matrices in terms of natural orbitals (NOrb). The diagonalization of 1-DM yields the NOrb's $\{\chi_i(\mathbf{1})\}_{i=1,K}$ and natural occupancies (NOcc's) $\{n_i\}_{i=1,K}$, whereas the eigenvectors and eigenvalues of the 2-DM are, respectively, the orthonormal 2-particle functions known as geminal orbitals $\{\Gamma_i(\mathbf{1},\mathbf{2})\}_{i=1,K}$ and the geminal occupancies or pair occupation numbers $\{\gamma_{ij}\}_{i,j=1,K}$.

Because the electronic Hamiltonian is a two-particle operator, as introduced in Eqs. 1.7 and 1.8, its expectation value can be expressed in terms of the 1-rDM and the 2-rDM,

$$E = \langle\widehat{H}\rangle = -\frac{1}{2}\int_{1\to1'}\nabla^2_{\mathbf{1}}\rho_1(\mathbf{1'};\mathbf{1})d\mathbf{1'} + \int\frac{\rho(\mathbf{1})}{|\boldsymbol{R}-\boldsymbol{r_1}|}d\mathbf{1} + \iint\frac{\rho_2(\mathbf{1'},\mathbf{1})}{|\boldsymbol{r'_1}-\boldsymbol{r_1}|}d\mathbf{1'}d\mathbf{1}, \quad (1.21)$$

and the 2-PD is, then, an important quantity to describe electron correlation, since V_{ee} is a functional of the 2-PD. [115]. T_e and V_{eN} are functionals of the 1-rDM;

particularly, V_{eN} is a functional of the diagonal elements of the 1-rDM, the electron density. The energy is, then, an exact functional of the 2-rDM [122,123]. This topic will be extended in section 1.2.3. The expectation value of the energy can then be obtained as

$$E = \sum_{ij} {}^1\mathrm{D}_{ij} h_{ij} + \frac{1}{2} \sum_{ijkl} {}^1\mathrm{D}_{ij}^{kl} g_{ijkl} \tag{1.22}$$

where h_{ij} and g_{ijkl} are the one- and two-electron integrals, related to the one- and two-electron operators defined in Eq. 1.8 [27].

Density matrices can be decomposed into spin contributions, that is, separating the elements containing the interactions between electron pairs with the same spin ($\alpha\alpha$) and with opposite spin ($\alpha\beta$), by integration over the spin coordinates. The spinless 2-rDM is formed by six contributions,

$$\begin{aligned}\rho_2(\boldsymbol{r_1}, \boldsymbol{r_2}; \boldsymbol{r'_1}, \boldsymbol{r'_2}) =& \rho_2^{\alpha\alpha\alpha\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2}; \boldsymbol{r'_1}, \boldsymbol{r'_2}) + \rho_2^{\beta\beta\beta\beta}(\boldsymbol{r_1}, \boldsymbol{r_2}; \boldsymbol{r'_1}, \boldsymbol{r'_2}) \\ &+ \rho_2^{\alpha\beta\alpha\beta}(\boldsymbol{r_1}, \boldsymbol{r_2}; \boldsymbol{r'_1}, \boldsymbol{r'_2}) + \rho_2^{\beta\alpha\beta\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2}; \boldsymbol{r'_1}, \boldsymbol{r'_2}) \\ &+ \rho_2^{\alpha\beta\beta\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2}; \boldsymbol{r'_1}, \boldsymbol{r'_2}) + \rho_2^{\beta\alpha\alpha\beta}(\boldsymbol{r_1}, \boldsymbol{r_2}; \boldsymbol{r'_1}, \boldsymbol{r'_2}),\end{aligned} \tag{1.23}$$

and its diagonal elements, the spinless 2-PD, reduce to

$$\rho_2(\boldsymbol{r_1}, \boldsymbol{r_2}) = \rho_2^{\alpha\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2}) + \rho_2^{\beta\beta}(\boldsymbol{r_1}, \boldsymbol{r_2}) + \rho_2^{\alpha\beta}(\boldsymbol{r_1}, \boldsymbol{r_2}) + \rho_2^{\beta\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2}). \tag{1.24}$$

Analogously, one can also decompose the 1-rDM into four components,

$$\rho_1(\boldsymbol{r_1}; \boldsymbol{r'_1}) = \rho_1^{\alpha\alpha}(\boldsymbol{r_1}; \boldsymbol{r'_1}) + \rho_1^{\beta\beta}(\boldsymbol{r_1}; \boldsymbol{r'_1}) + \rho_1^{\alpha\beta}(\boldsymbol{r_1}; \boldsymbol{r'_1}) + \rho_1^{\beta\alpha}(\boldsymbol{r_1}; \boldsymbol{r'_1}). \tag{1.25}$$

and the electron density constitutes of elements $\rho_1^{\alpha}(\boldsymbol{r_1})$ and $\rho_1^{\beta}(\boldsymbol{r_1})$. Working with spinless density matrices becomes opportune and suitable in quantum chemistry, since they are simply spin-traced versions of the complete matrix.

The traces of the spin components of the 2-PD are

$$\mathrm{Tr}\left[\rho_2^{\alpha\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2})\right] = \int \rho_2^{\alpha\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2}) d\boldsymbol{r_1} d\boldsymbol{r_2} = N^{\alpha}(N^{\alpha} - 1) \tag{1.26}$$

$$\mathrm{Tr}\left[\rho_2^{\alpha\beta}(\boldsymbol{r_1}, \boldsymbol{r_2})\right] = \int \rho_2^{\alpha\beta}(\boldsymbol{r_1}, \boldsymbol{r_2}) d\boldsymbol{r_1} d\boldsymbol{r_2} = N^{\alpha} N^{\beta}, \tag{1.27}$$

and $\rho_2^{\beta\beta}$ and $\rho_2^{\beta\alpha}$ take analogous expressions. In the particular case of a closed-shell electronic system, Eqs. 1.26 and 1.27 reduce to

$$\mathrm{Tr}\left[\rho_2^{\alpha\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2})\right] = \int \rho_2^{\alpha\alpha}(\boldsymbol{r_1}, \boldsymbol{r_2}) d\boldsymbol{r_1} d\boldsymbol{r_2} = \frac{N(N-2)}{4} \tag{1.28}$$

and

$$\mathrm{Tr}\left[\rho_2^{\alpha\beta}(\boldsymbol{r_1},\boldsymbol{r_2})\right] = \int \rho_2^{\alpha\beta}(\boldsymbol{r_1},\boldsymbol{r_2})d\boldsymbol{r_1}d\boldsymbol{r_2} = \frac{N^2}{4}. \tag{1.29}$$

The trace of the complete, spinless 2-PD is obtained by accounting for all the spin components, $\mathrm{Tr}[\rho_2(\boldsymbol{r_1},\boldsymbol{r_2})] = \mathrm{Tr}[\rho_2^{\alpha\alpha}(\boldsymbol{r_1},\boldsymbol{r_2})] + \mathrm{Tr}[\rho_2^{\beta\beta}(\boldsymbol{r_1},\boldsymbol{r_2})] + \mathrm{Tr}[\rho_2^{\alpha\beta}(\boldsymbol{r_1},\boldsymbol{r_2})] + \mathrm{Tr}[\rho_2^{\beta\alpha}(\boldsymbol{r_1},\boldsymbol{r_2})]$. The case of the 1-electron density reduces to very simple quantities. Since its trace corresponds to the number of electrons, the spin counterparts are normalized to the number of alpha and beta electrons separately.

The 2-PD can be separated in the same manner as the E_c definition (Eq. 1.1), being the sum of the uncorrelated HF 2-PD and a 2-PD that contains all the remaining correlation in the system, which we name the correlation pair density, $\Delta\rho_2^\mathrm{c}$,

$$\rho_2(\boldsymbol{r_1},\boldsymbol{r_2}) = \rho_2^\mathrm{HF}(\boldsymbol{r_1},\boldsymbol{r_2}) + \Delta\rho_2^\mathrm{c}(\boldsymbol{r_1},\boldsymbol{r_2}), \tag{1.30}$$

and an analogous separation can be made within the framework of KS-DFT,

$$\rho_2(\boldsymbol{r_1},\boldsymbol{r_2}) = \rho_2^\mathrm{KS}(\boldsymbol{r_1},\boldsymbol{r_2}) + \rho_2^\mathrm{c,KS}(\boldsymbol{r_1},\boldsymbol{r_2}). \tag{1.31}$$

The correlation 2-PD $\Delta\rho_2^\mathrm{c}$ is also known as the cumulant of the 2-PD, usually labeled as $\lambda(\mathbf{1},\mathbf{2})$ in the literature, and also ${}^2\Gamma_{ij}^{kl}$ if expressed in the orbital representation [124,125]. The cumulant term is the part of the 2-rDM that cannot be described in terms of the 1-rDM [124] and is considered to contain all the correlation effects that the SD approximation is not able to describe. Considering the orbital expansion of the 2-rDM, the 2-DM,

$${}^2\mathrm{D}_{ij}^{kl} = 2\,{}^1\mathrm{D}_i^k \wedge {}^1\mathrm{D}_j^l + {}^2\Gamma_{ij}^{kl} = {}^1\mathrm{D}_i^k\,{}^1\mathrm{D}_j^l - {}^1\mathrm{D}_i^l\,{}^1\mathrm{D}_j^k + {}^2\Gamma_{ij}^{kl}, \tag{1.32}$$

where the 2-DM is expressed in terms of the orbital representation of the cumulant matrix, and the Grassmann wedge product of the 1-DM [126,127].

The Coulomb and exchange interactions can be described by a unique quantity, the exchange-correlation density [128],

$$\rho_2(\boldsymbol{r_1},\boldsymbol{r_2}) = \rho(\boldsymbol{r_1})\rho(\boldsymbol{r_2}) - \rho_2^\mathrm{xc}(\boldsymbol{r_1},\boldsymbol{r_2}). \tag{1.33}$$

Note that the exchange-correlation density integrates to $-N$. It is plain to see that electrons are statistically uncorrelated if and only if $\rho_2^\mathrm{xc}(\boldsymbol{r_1},\boldsymbol{r_2}) = 0$. The exchange-correlation density is important for the definition of McWeeny's correlation hole (see Chapter 1.3.1) and is sometimes used as the main ingredient in designing bonding analysis tools, such as in the electron delocalization index (DI) [129].

1.2.2 The Hartree–Fock approximation and its density matrices

Even though some correlation facts about the HF method have already been introduced, in this section some concepts are summarized and emphasized. Prior to the HF method, the Hartree Product was one of the first approaches to construct a wavefunction. It is one of the simplest ways of constructing a wavefunction, consisting of the product of N one-electron functions or orbitals [130, 131],

$$\Psi^{\mathrm{H}}(\mathbf{1},\ldots,\boldsymbol{N}) = \psi_1(\mathbf{1})\ldots\psi_1(\boldsymbol{N}). \tag{1.34}$$

The Hartree Product, however, does not consider the fermionic nature of electrons, for which it is completely uncorrelated; that is, the wavefunction is not antisymmetrized, thus the Pauli principle is not fulfilled. Electrons are treated as bosons and do not interact at all. This is reflected in the 2-PD of the Hartree wavefunction, formed by the product of individual electron densities for any spin case, $\rho_2^{\mathrm{H}}(\mathbf{1},\mathbf{2}) = \rho^{\mathrm{H}}(\mathbf{1})\rho^{\mathrm{H}}(\mathbf{2})$. The inclusion of Slater determinants in the wavefunction description causes the Pauli exclusion principle to be fulfilled,

$$\Psi^{\mathrm{HF}}(\mathbf{1},\ldots,\boldsymbol{N}) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(\mathbf{1}) & \cdots & \psi_1(\boldsymbol{N}) \\ \vdots & \ddots & \vdots \\ \psi_N(\mathbf{1}) & \cdots & \psi_N(\boldsymbol{N}) \end{vmatrix}, \tag{1.35}$$

leading electron pairs with the same spin to be correlated, as the interchange of any pair of electron coordinates implies a change in the sign of the wavefunction due to determinant properties,

$$\Psi^{\mathrm{HF}}(\ldots,\boldsymbol{i},\ldots,\boldsymbol{j},\ldots) = -\Psi^{\mathrm{HF}}(\ldots,\boldsymbol{j},\ldots,\boldsymbol{i},\ldots). \tag{1.36}$$

This makes the HF approximation a method where the electron motions are no longer completely independent. However, whereas exchange interactions are now present, Coulombic interactions are still not contemplated, which leaves the motion of two electrons with unlike spins to be uncorrelated. Therefore, in HF (and for any one-determinant wavefunction), only electron pairs with like-spin are partially correlated. The HF method employs an approximated repulsive field generated by the electron cloud in the system, where the repulsion felt by an electron is an average potential created by the rest of electrons (mean-field approximation). This underestimation of electron repulsion causes electron pairs to be closer. The HF 2-PD illustrates the partially correlated behavior of electrons,

$$\rho_2^{\mathrm{HF}}(\mathbf{1},\mathbf{2}) = \rho^{\mathrm{HF}}(\mathbf{1})\rho^{\mathrm{HF}}(\mathbf{2}) - \rho_1^{\mathrm{HF}}(\mathbf{1};\mathbf{2})\rho_1^{\mathrm{HF}}(\mathbf{2};\mathbf{1}). \tag{1.37}$$

The product of individual electron densities depicts the Coulomb interactions and, consequently, reflects the lack of correlation between electrons in position and spin 1 and 2. The second term, described by the HF 1-rDM, represents the exchange interactions [25]. The 1-rDM product in Eq. 1.37 can be equivalently expressed as $|\rho_1^{\mathrm{HF}}(\mathbf{1};\mathbf{2})|^2$.

One can split Eq. 1.37 into spin contributions. The HF 2-PD that contains the opposite-spin interactions ($\alpha\beta$) is composed by an independent product of individual one-electron densities,

$$\rho_2^{\mathrm{HF},\alpha\beta}(\boldsymbol{r_1},\boldsymbol{r_2}) = \rho^{\mathrm{HF},\alpha}(\boldsymbol{r_1})\rho^{\mathrm{HF},\beta}(\boldsymbol{r_2}), \tag{1.38}$$

and the counterpart describing the same-spin interactions ($\alpha\alpha$) reads

$$\rho_2^{\mathrm{HF},\alpha\alpha}(\boldsymbol{r_1},\boldsymbol{r_2}) = \rho^{\mathrm{HF},\alpha}(\boldsymbol{r_1})\rho^{\mathrm{HF},\alpha}(\boldsymbol{r_2}) - |\rho_1^{\mathrm{HF},\alpha\alpha}(\boldsymbol{r_1};\boldsymbol{r_2})|^2. \tag{1.39}$$

In other words, HF does not correlate electrons of different spins, which are treated as statistically independent quantities.

When the set of orbitals chosen is the NOrb basis (Eq. 1.20), the orbital representation of the HF 2-PD per each electron-pair spin component takes the form of

$${}^2\mathrm{D}_{ij,kl}^{\mathrm{HF},\alpha\beta} = n_i^{\mathrm{HF},\alpha} n_j^{\mathrm{HF},\beta} \delta_{ik}\delta_{jl} \tag{1.40}$$

and

$${}^2\mathrm{D}_{ij,kl}^{\mathrm{HF},\alpha\alpha} = n_i^{\mathrm{HF},\alpha} n_j^{\mathrm{HF},\alpha} \left(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}\right) \tag{1.41}$$

where $\left\{n_i^{\mathrm{HF}}\right\}_{i=1,K}$ are the HF occupations, which take values of either 1 or 0, and the Kronecker deltas define the Coulombic ($\delta_{ik}\delta_{jl}$) and exchange ($\delta_{il}\delta_{jk}$) terms.

At this point, it is opportune to introduce the single-determinant (SD) ansatz to the 2-PD [121]. The SD 2-PD takes the same expression as the HF 2-PD (Eq. 1.37), but uses a generic 1-rDM instead:

$$\rho_2^{\mathrm{SD}}\left(\rho_1,\mathbf{1},\mathbf{2}\right) = \rho(\mathbf{1})\rho(\mathbf{2}) - |\rho_1(\mathbf{1};\mathbf{2})|^2, \tag{1.42}$$

where, compared to Eq. 1.37, the density elements in Eq. 1.42 are not the HF densities. This expression recovers the HF 2-PD when the HF 1-rDM is used,

$$\rho_2^{\mathrm{HF}}(\mathbf{1},\mathbf{2}) \equiv \rho_2^{\mathrm{SD}}\left(\rho_1^{\mathrm{HF}},\mathbf{1},\mathbf{2}\right) = \rho^{\mathrm{HF}}(\mathbf{1})\rho^{\mathrm{HF}}(\mathbf{2}) - |\rho_1^{\mathrm{HF}}(\mathbf{1};\mathbf{2})|^2. \tag{1.43}$$

The orbital expansion of the SD 2-PD is also described by Eqs. 1.40 and 1.41. Note that, since the exact 1-rDM is used instead of the HF 1-rDM, the occupations taken in these expressions are fractionary. As a consequence, the SD 2-PD

is not normalized by the number of electron pairs, but with a larger quantity, $\mathrm{Tr}\left[\rho_2^{\mathrm{SD}}(\mathbf{1},\mathbf{2})\right] > N(N-1)$. This is caused by the $\alpha\alpha$ elements in the SD 2-PD. The normalization coefficient for the $\alpha\beta$ component corresponds to the one depicted in Eq. 1.29, or in Eq. 1.27 for open-shell cases, but the trace for the $\alpha\alpha$ 2-PD is larger than the normalization coefficient expressed in Eq. 1.28 (or Eq. 1.26 for open-shell cases),

$$\mathrm{Tr}\left[\rho_2^{\mathrm{SD},\alpha\beta}(\mathbf{1},\mathbf{2})\right] = \int \rho_2^{\mathrm{SD},\alpha\beta}(\boldsymbol{r_1},\boldsymbol{r_2})d\boldsymbol{r_1}d\boldsymbol{r_2} = \frac{N^2}{4} \tag{1.44}$$

$$\mathrm{Tr}\left[\rho_2^{\mathrm{SD},\alpha\alpha}(\mathbf{1},\mathbf{2})\right] = \int \rho_2^{\mathrm{SD},\alpha\alpha}(\boldsymbol{r_1},\boldsymbol{r_2})d\boldsymbol{r_1}d\boldsymbol{r_2} > \frac{N(N-2)}{4}. \tag{1.45}$$

Also as a consequence, the energy obtained from the SD 2-DM will be larger (and, therefore, variational) than the HF energy [132].

1.2.3 Approximate density matrices

The knowledge of the 2-rDM of the system permits the evaluation of the Hamiltonian and obtaining its expectation value by following Eq. 1.21. It feels only natural, then, to solve the Schrödinger equation employing 2-rDMs alone and avoiding the use of the wavefunction, the dimension of which implies a huge demand of computational capacity. Then, one could variationally optimize a trial 2-rDM instead of a wavefunction, until reaching the minimum energy. This approach is treated in reduced density matrix functional theory (rDMFT) [133–135] and it was first introduced by Gilbert in 1975 as an extension to the Hohenberg-Kohn theorem for nonlocal external potentials [133].

A brief introduction to rDMFT is given in the forthcoming sections. In Chapter 5, a benchmark of a set of reduced density matrix functional approximations (rDMFA) is presented. The study pretends to complement a former benchmark, where the approximations were tested with a physical model [136]. The current benchmark study presented in Chapter 5 studies similar properties in a set of different molecules. The assessed approximations are presented after the introduction to the theory in section 1.2.3.3.

1.2.3.1 Fundamentals: Gilbert's theorem

In the framework of DFT, Hohenberg and Kohn proved a one-to-one mapping between the ground state wavefunction and the ground state electron density [137]. Because of this, the electronic energy can be expressed as a functional of the electron density. Analogously, Gilbert proved an extension of the Hohenberg and Kohn

theorem, where a one-to-one mapping between the ground state wavefunction and the ground state 1-rDM is established [133],

$$\Psi^{\text{gs}}(\mathbf{1},\ldots,\boldsymbol{N}) \Leftrightarrow \rho_1^{\text{gs}}(\mathbf{1};\mathbf{1'}) \Rightarrow E^{\text{gs}}. \tag{1.46}$$

The derivation of Gilbert's theorem is very similar to the Hohenberg-Kohn theorem. Consider two distinc Hamiltonians $\hat{H}$ and $\hat{H}'$ that are defined by two different nonlocal external potentials $\hat{V}_{eN} = v(\mathbf{1},\mathbf{1'})$ and $\hat{V}'_{eN} = v'(\mathbf{1},\mathbf{1'})$. Assuming that the corresponding ground state wavefunctions Ψ^{gs} and $\Psi^{\text{gs'}}$ are non-degenerate and different, the variational principle states that

$$\langle\Psi^{\text{gs'}}|\hat{H}|\Psi^{\text{gs'}}\rangle - \langle\Psi^{\text{gs}}|\hat{H}|\Psi^{\text{gs}}\rangle = E^{\text{gs'}} + \langle\Psi^{\text{gs'}}|\hat{V}_{eN} - \hat{V}'_{eN}|\Psi^{\text{gs'}}\rangle - E^{\text{gs}} > 0 \tag{1.47}$$

and

$$\langle\Psi^{\text{gs}}|\hat{H}'|\Psi^{\text{gs}}\rangle - \langle\Psi^{\text{gs'}}|\hat{H}'|\Psi^{\text{gs'}}\rangle = E^{\text{gs}} + \langle\Psi^{\text{gs}}|\hat{V}'_{eN} - \hat{V}_{eN}|\Psi^{\text{gs}}\rangle - E^{\text{gs'}} > 0, \tag{1.48}$$

where E^{gs} and $E^{\text{gs'}}$ are the corresponding ground state energies. If the difference of external potentials is $v(\mathbf{1},\mathbf{1'}) - v'(\mathbf{1},\mathbf{1'}) = \delta v(\mathbf{1},\mathbf{1'})$, the addition of Eqs. 1.47 and 1.48 yields

$$\begin{aligned}\langle\Psi^{\text{gs'}}|\hat{V}_{eN} - \hat{V}'_{eN}|\Psi^{\text{gs'}}\rangle + \langle\Psi^{\text{gs}}|\hat{V}'_{eN} - \hat{V}_{eN}|\Psi^{\text{gs}}\rangle = \\ \int \delta v(\mathbf{1},\mathbf{1'})\rho_1^{\text{gs'}}(\mathbf{1};\mathbf{1'})d\mathbf{1}d\mathbf{1'} - \int \delta v(\mathbf{1},\mathbf{1'})\rho_1^{\text{gs}}(\mathbf{1};\mathbf{1'})d\mathbf{1}d\mathbf{1'} = \\ \int \delta v(\mathbf{1},\mathbf{1'})\left[\rho_1^{\text{gs'}}(\mathbf{1},\mathbf{1'}) - \rho_1^{\text{gs}}(\mathbf{1},\mathbf{1'})\right]d\mathbf{1}d\mathbf{1'} > 0\end{aligned} \tag{1.49}$$

Hence, the existence of an energy functional of the 1-rDM is proven because the difference of 1-rDMs in Eq. 1.49 must be nonzero. rDMFT, then, takes the 1-rDM as the main quantity for building approximations. If the approximations are built in the NOrb representation of the 1-rDM, that is, by means of orbital occupancies, the theory is named natural orbital functional theory or NOFT [138].

T_e and V_{eN} can be determined exactly because they are explicit functionals of the 1-rDM, as introduced in Eq. 1.21. Instead, because V_{ee} is a functional of the 2-PD, it needs to be approximated in terms of the 1-rDM:

$$\begin{aligned}E[\rho_1(\mathbf{1},\mathbf{1'})] =& V_{ee}[\rho_1(\mathbf{1},\mathbf{1'})] + T_e[\rho_1(\mathbf{1};\mathbf{1'})] + V_{eN}[\rho_1(\mathbf{1};\mathbf{1'})] \\ =& V_{ee}[\rho_1(\mathbf{1},\mathbf{1'})] + \int_{1\to 1'}\left[-\frac{1}{2}\nabla_{1'}\nabla_1 + \frac{1}{|\boldsymbol{R}-\boldsymbol{r_1}|}\right]\rho_1(\mathbf{1};\mathbf{1'})d\mathbf{1}\end{aligned} \tag{1.50}$$

Approximating V_{ee} in terms of the 1-rDM permits lower storage and data manipulation in computations. Considering the natural orbital representation of the matrices,

an explicit calculation of V_{ee} involves the handling of K^4 terms from the 2-DM, being K the size of the orbital basis. Instead, 1-DMs are composed of K^2 elements, which makes them more appropriate objects for computations.

1.2.3.2 The N-representability problem

It has been introduced in section 1.2.1 that the rank reduction of a density matrix implies a loss of information that cannot be recovered, being thus far impossible to restore a higher-order density matrix from a lower-order one. It is, then, sensible to impose physical requirements for the reconstruction of approximate 2-rDMs in terms of 1-rDMs; otherwise, using an unphysical 2-rDM in the Hamiltonian may lead to non-variational energies [120].

Some trivial restrictions can be imposed, as forcing that the sum of the trace elements of the 2-DM must result in the total number of electron pairs (that is, the 2-DM must be appropriately normalized, Eqs. 1.28 and 1.29). This imposition is known as the sum rule. However, there exist a set of requirements that a density matrix must fulfill. This set of requirements is referred as the N-representability conditions, and guarantees that the constructed 2-rDM emanates from a fermionic N-wavefunction [120, 123]; in other words, the approximate 2-rDM must be obtainable from the integration of a fermionic wavefunction as described in Eq. 1.16. Otherwise, a non N-representable 2-rDM can lead to energies below the exact result from variational optimizations [120]. However, up to date it does not exist a direct way to ensure the N-representability of the 2-rDM without going back up to the original N-DM, which is not efficient for computations. The concern and assessment of such an issue is what is known as the N-representability problem [120]. Note that the N-representability problem affects not only the 2-rDM but any approximated n-rDM and their associated n-densities. An rDM that can be mapped back to an antisymmetric N-wavefunction is called to be N-representable.

Even though the complete set of conditions are not completely known for the n-rDMs, the conditions needed to ensure the N-representability for the 1-rDM and 2-rDM have been widely studied [139–141]. For the particular case of the 1-rDM, the full set of sufficient conditions is known. Considering the NOrb representation

of the 1-rDM, the 1-DM, the conditions are [139]:

$$\sum_{i}^{K} n_i = N \tag{1.51}$$

$$0 \leq n_i \leq 1 \tag{1.52}$$

That is, the sum of NOcc must be the total number of electrons in the system N, and the occupation numbers must range between zero and one.

The whole set of sufficient conditions for the 2-rDM is known, yet their imposition is unfeasible since the rest of higher-order density matrices are needed (that is, from 3-rDM to N-rDM) [142], leading to an impractical assessment of the approach. Instead, a group of necessary (yet not sufficient) conditions is imposed. They are probability-based and, being probabilities, concern the evaluation of the non-negative character of the 2-rDM elements. These conditions are constructed in matrix form, and their eigenvalues must be positive semidefinite for a 2-rDM to be N-representable. Because of this, this group of conditions is called positivity conditions. A set of positivity conditions exists for any n-rDM [139–141].

The positivity condition matrices are defined from an operator $\hat{A}$ composed of strings of annihilation and creation operators, $\hat{a}_{i\sigma}$ and $\hat{a}^{\dagger}_{i\sigma}$ respectively, in second quantization terminology [27]. The positivity conditions require the expectation value from this operator to be positive semidefinite [143]:

$$B_{i,j} = \langle \Psi | \hat{A}_i^{\dagger} \hat{A}_j | \Psi \rangle \geq 0, \tag{1.53}$$

where $B_{i,j}$ is the expectation value. For the 2-rDM, operator $\hat{A}$ is the product of two creation and annihilation operators [144]:

$$\hat{A} \in \left\{ \hat{a}_i^{\dagger}\hat{a}_j^{\dagger}, \hat{a}_i^{\dagger}\hat{a}_j, \hat{a}_i\hat{a}_j \right\} \tag{1.54}$$

Three matrices result of doing all the possible combinations of $\hat{A}$ elements from Eq. 1.54 into Eq. 1.53, and they are known as the P, Q and G matrices [119, 139, 141, 145, 146]:

$$P^{\sigma\sigma'}_{ij,kl} = \langle \Psi | a^{\dagger}_{i\sigma} a^{\dagger}_{j\sigma'} a_{k\sigma} a_{l\sigma'} | \Psi \rangle \tag{1.55}$$

$$Q^{\sigma\sigma'}_{ij,kl} = \langle \Psi | a_{i\sigma} a_{j\sigma'} a^{\dagger}_{k\sigma} a^{\dagger}_{l\sigma'} | \Psi \rangle \tag{1.56}$$

$$G^{\sigma\sigma'}_{ij,kl} = \langle \Psi | a_{i\sigma} a^{\dagger}_{j\sigma'} a^{\dagger}_{k\sigma} a_{l\sigma'} | \Psi \rangle \tag{1.57}$$

These matrices exist for any spin-combination case, since $\sigma\sigma' = \alpha\alpha, \alpha\beta, \beta\alpha, \beta\beta$ [143]. The diagonal elements of these P, Q and G matrices are related to different probabilities of electron distributions among orbitals. Therefore, the diagonal elements of

these matrices must be non-negative, or positive semidefinite, for *any* orbital representation. The positive conditions can be evaluated through the eigenvalues of the P, Q and G matrices [146] which, in the NOrb representation, results in restrictions to be applied to the NOccs. The definition of the P matrix actually corresponds to the definition of the 2-DM in second quantization language, so the P condition forces the geminal occupancies to be positive semidefinite. On the other hand, the Q and G conditions indicate the positive semidefinite character of the holes $(1-n)$ and the particle-hole $n(1-n)$ probabilities, respectively. Fig. 1.6 contains a representation summary of the positivity conditions. Note that similar matrices can also be obtained for any n-rDM.

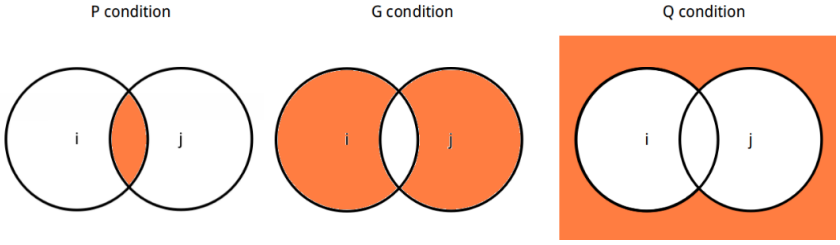


Figure 1.6: Venn diagrams representing the probabilities of occupying orbitals i and j behind P (particle-particle), G (particle-hole) and Q (hole-hole) conditions for the 2-DM.

1.2.3.3 Reduced density matrix functional theory approximations

The rDMFT approximations examined in this thesis consider the Coulomb (J), exchange (K), and time-inversion (L) two-electron integrals in the orbital representation [147, 148]. Despite the vast majority of rDMFAs are defined in the orbital representation and employ NOccs, there are other rDMFAs that use the space representation to develop the approximations (see, for instance, the ΔNO functional [52, 149, 150]). According to the approach considered for their construction, rDMFAs can be classified into two groups: the K-functionals and the PNOFs. The functionals presented hereafter are the ones that are studied in Chapter 5.

1.2.3.3.1 K-functionals

K-functionals are approximations that use the HF expression for the Coulombic terms (Eqs. 1.41 and 1.40) and propose an expression to approximate the exchange elements in the 2-DM. The electrostatic interactions are left untouched as some authors suggested that correlation effects can be approximated with just exchange

integrals [151, 152]. In the literature, this group of rDMFAs is referred to as JK-functionals because the approximate 2-DM elements are the ones accompanying either the Coulomb and the exchange integrals. In this thesis, we consider that it is more appropriate to name them K-functionals alone since the approximations only affect the exchange terms. K-functionals use the following expressions to construct 2-DM approximations [146]:

$$
\begin{aligned}
{}^2\mathrm{D}^{X,\alpha\beta}_{ij,kl} &= n_i^\alpha n_j^\beta \delta_{ik}\delta_{jl}, && (1.58)\\
{}^2\mathrm{D}^{X,\alpha\alpha}_{ij,kl} &= n_i^\alpha n_j^\alpha \delta_{ik}\delta_{jl} - f^X(n_i^\alpha, n_j^\alpha)\delta_{il}\delta_{jk}, && (1.59)
\end{aligned}
$$

where $f^X(n_i^\alpha, n_j^\alpha)$ (hereafter expressed as $f^X(n_i, n_j)$ for simplicity) is the approximate expression proposed by functional "X" for the exchange terms. Since the derivation of Gilbert's theorem, a plethora of K-functionals are present in the literature. In Chapter 5, a group of eight K-functional approximations have been benchmarked. The $f^X(n_i, n_j)$ expressions taken by each functional are gathered in Table 1.1.

Not specifically designed to be an rDMFA, the SD ansatz is the most direct approach in approximating the 2-DM (Eq. 1.42), taking the HF expression (Eq. 1.37) for describing both Coulombic and exchange terms [121]. Because the HF expression is used but not the HF occupancies, the SD 2-DM is not purely N-representable and, for instance, does not follow the sum rule ($\mathrm{Tr}\left[{}^2\mathrm{D}^{\mathrm{SD}}_{ij,kl}\right] \equiv \mathrm{Tr}\left[\rho_2^{\mathrm{SD}}(\boldsymbol{r_1}, \boldsymbol{r_2})\right] \neq N$). The SD 2-DM coincides with the HF 2-DM when canonical orbitals and occupations are used, and optimization of occupancies and orbitals using the 2-DM SD leads to the HF energy, occupancies, and orbitals.

The first actual approximation of the 2-DM in the rDMFT framework was derived independently by Müller by one side, and by Buijse and Bearends by another one. Some authors call it the "Corrected Hartree" functional [153], but it is best known by the acronym of the author's initials, MBB [151,152,154]. The approximation considered the requirement of producing a minimal amount of deviations of the Pauli principle condition, and was derived from Coulomb and Fermi holes. It provides a qualitatively correct correlation description, as in the H_2 dissociation energy curve and the He isoelectronic series, when NOrb's and NOcc's are not treated self-consistently [151, 152, 155]. On the other hand, Goedecker and Umrigar designed the GU approximation to correct the self-interacting error in MBB by omitting all diagonal Coulomb and exchange integrals [156]. The rest of the elements take the same expression as in MBB. With this correction, the total energies for atoms and molecules at equilibrium geometries improve with respect to MBB, yet it yields

worse dissociation descriptions [146,155,157]. A more physically motivated approach is the one proposed by Gritsenko and coworkers, who designed the BBC1, BBC2, and BBC3 approximations to improve the MBB functional [155]. In this thesis, the BBC2 approximation has been studied. This rDMFA introduces the definition of the Fermi level F_L (being the last orbital whose occupancy is greater or equal to $N/2$) to define the orbital occupancies and the interactions between electron pairs change according to the orbitals in which are found. The interaction between highly- (below F_L) and weakly-occupied (above F_L) orbitals are described with the MBB expression, and the HF expression is recovered in the case of two orbitals found below the F_L. Other cases are summarized in Table 1.1. MBB and BBC2 are designed to fulfill the sum rule, but GU does not because of the straightforward correction introduced.

On the other hand, we find the CA and CGA functionals, which were proposed mainly by Csányi and Arias [153,158]. They proposed a tensor product expansion of the 2-rDM to construct the approximation and the approximations were designed to reproduce the homogeneous electron gas. CGA was later presented as an improvement of CA for higher densities, a regime in which CA showed deficiencies [158]. A difference with respect to the MBB group of functionals is that the CA and CGA expressions consider the particle-hole symmetry $n(1-n)$, as it is shown in Table 1.1. According to the authors, the CGA functional is comparable to the generalized-gradient approximation performance in atoms [158]. Both approaches recover the normalization coefficient of the 2-DM.

The latter group of K-functional approximations considered in this thesis is the one derived by Marques and Lathiotakis, the ML and MLSIC functionals [159]. The square root in MBB is replaced with a Padé approximant, and include three parameters that were empirically fitted to reproduce the correlation energies of the G2 set of molecules. Because of this fitting, the sum rule is not satisfied in either of both cases. MLSIC is the self-interaction-corrected version of ML and, therefore, ML presents a unique expression for all the exchange interactions, whereas MLSIC reoptimizes the fitting parameters after omitting all diagonal Coulomb and exchange elements to avoid the double interaction counting.

Functional	$f^X(n_i, n_j)$		
SD	$n_i n_j$	$\forall i, j$	
MBB	$(n_i n_j)^{1/2}$	$\forall i, j$	
GU	$(n_i n_j)^{1/2}$	$i \neq j$	
	$n_i n_j$	$i = j$	
BBC2	n_i	$i = j$	
	$n_i n_j$	$i \neq j \wedge i \in [1; F_L] \wedge j \in [1; F_L]$	
	$-(n_i n_j)^{1/2}$	$i \neq j \wedge i \in (F_L; \infty) \wedge j \in (F_L; \infty)$	
	$(n_i n_j)^{1/2}$	otherwise	
CA	$[(n_i(1-n_i)n_j(1-n_j)]^{1/2} + n_i n_j$	$\forall i, j$	
CGA	$\frac{1}{2}\left[[(n_i(2-n_i)n_j(2-n_j)]^{1/2} + n_i n_j\right]$	$\forall i, j$	
ML	$n_i n_j(a + b\, n_i n_j)(1 + c\, n_i n_j)$	$\forall i, j$	$a = 126.3101$ $b = 2213.33$ $c = 2338.64$
MLSIC	$n_i n_j(a + b\, n_i n_j)/(1 + c\, n_i n_j)$	$i \neq j$	$a = 1298.78$
	$n_i n_j$	$i = j$	$b = 35114.4$ $c = 36412.2$

Table 1.1: Exchange expressions $f^X(n_i, n_j)$ (Eq. 1.59) per each K-functional studied in this thesis.

1.2.3.3.2 PNOF functionals

The second group in DMFT approximations are the ones that are based on the

spectral expansion of the cumulant of the 2-DM (Eq. 1.32) [160]. As introduced in section 1.2.1, the cumulant of the 2-DM is defined as the part of the 2-DM that includes the electron correlation information that is not included in the HF method (that is, correlation effects beyond Fermi correlation) [124, 125],

$$ {}^2D^{\alpha\beta}_{ij,kl} = n_i^\alpha n_j^\beta \delta_{ik}\delta_{jl} + {}^2\Gamma^{\alpha\beta}_{ij,kl}, \tag{1.60} $$

$$ {}^2D^{\alpha\alpha}_{ij,kl} = n_i^\alpha n_j^\alpha (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) + {}^2\Gamma^{\alpha\alpha}_{ij,kl}. \tag{1.61} $$

The cumulant matrices contain the missing correlation effects that the SD approximation of the 2-DM does not include. Because the expressions of the cumulant matrices are not known, an approach to reconstruct the correlated 2-DM consists in providing approximations to the same- and opposite-spin cumulant matrices. The PNOF functionals (after "Piris Natural Orbital Functional") approximate the antisymmetric cumulant matrices by splitting them into two auxiliary matrices known as Δ and Π [160],

$$ {}^2\Gamma^{\alpha\beta}_{ij,kl} = -\Delta^{\alpha\beta}_{ij}\delta_{ik}\delta_{jl} + \Pi_{ik}\delta_{ij}\delta_{kl}, \tag{1.62} $$

$$ {}^2\Gamma^{\alpha\alpha}_{ij,kl} = -\Delta^{\alpha\alpha}_{ij}\left(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}\right). \tag{1.63} $$

Note that these auxiliary matrices are composed of two indices in contrast to the cumulant matrix, composed of four indices. This is done to simplify the cumulant. Another difference with recpect to the K-functionals is that PNOFs also provide approximations for the Coulomb elements, and include Coulomb (J), exchange (K) and the exchange and time-inversion (L) integrals [160].

The different expressions given for both matrices produce different versions of the PNOFs, named PNOFi with $i = 2, \ldots, 7$ [55, 148, 161–166]. In general, the functionals are constructed to respect the N-representability conditions and the sum rule. They were not designed to consider open-shell cases, yet this issue is currently being explored [167]. The particle-hole symmetries are considered in most of the expressions for the functionals. For the Δ matrix, the same expression is used for both spin cases, except in PNOF3 [162]. Table 1.2 gathers the expressions for the Δ and Π matrices.

PNOF5 and PNOF6 were initially designed to follow a perfect pairing approach, which forces the coupling of occupancies of different orbitals, $n_i + n_j = 1$ [148, 164–166]. n_i will be coupled to n_j when one of these orbitals is found below the F_L and the other one above it. An extended approach was later on considered for PNOF5,

where the occupancy of the orbital below the F_L is coupled to the occupancies of a set of orbitals above the F_L. This extension was named PNOF5e. In this thesis, we have also considered this extended approach in PNOF6 (PNOF6e).

PNOF6 is defined in terms of a variable S_γ^x [148] and can take different definitions if the orbitals considered are placed above (up, $x = \mathrm{u}$) or below (down, $x = \mathrm{d}$) the Fermi level. An average of both cases can also be considered (half, $x = \mathrm{h}$). The formulas for the PNOF6x functional are summarized in Table 1.2.

In Chapter 5, the PNOF and K-functionals presented are benchmarked using a set of molecules in order to test the soundness of their expressions, and some chemical properties. The approach followed to do the study does not allow the analysis PNOF5e and PNOF6e as presented in this section due to the perfect-pairing constraints imposed by the functionals. Instead, the restriction-free versions of these functionals are employed.

Functional	Δ_{ij}	Π_{ij}	
PNOF2	$h_i h_j$	$\sqrt{n_i n_j} + \sqrt{h_i h_j} + T_{ij}$	$i \wedge j \in [1, F_L]$
	$h_i n_j \left(\frac{1-s_F}{s_F}\right)$	$\sqrt{n_i n_j} - \sqrt{h_i n_j} + T_{ij}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i h_j \left(\frac{1-s_F}{s_F}\right)$	$\sqrt{n_i n_j} - \sqrt{n_i h_j} + T_{ij}$	$i \in (F_L, K] \wedge j \in [1, F_L]$
	$n_i n_j$	T_{ij}	$i \wedge j \in (F_L, K]$
PNOF3	$h_i h_j$	$n_i n_j - \sqrt{n_i n_j}$	$i \wedge j \in [1, F_L]$
	$h_i n_j \left(\frac{1-s_F}{s_F}\right)$	$n_i n_j - \sqrt{n_i n_j} - \sqrt{h_i n_j}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i h_j \left(\frac{1-s_F}{s_F}\right)$	$n_i n_j - \sqrt{n_i n_j} - \sqrt{n_i h_j}$	$j \in [1, F_L] \wedge i \in (F_L, K]$
	$n_i n_j$	$n_i n_j + \sqrt{n_i n_j}$	$i \wedge j \in (F_L, K]$
PNOF4	$h_i h_j)$	$-\sqrt{h_i h_j}$	$i \wedge j \in [1, F_L]$
	$h_i n_j \left(\frac{1-s_F}{s_F}\right)$	$-\sqrt{\left(\frac{h_i n_j}{S_F}\right)\left(n_i - n_j + \frac{h_i n_j}{S_F}\right)}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i h_j \left(\frac{1-s_F}{s_F}\right)$	$-\sqrt{\left(\frac{n_i h_j}{S_F}\right)\left(-n_i + n_j + \frac{n_i h_j}{S_F}\right)}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i n_j$	$\sqrt{n_i n_j}$	$i \wedge j \in (F_L, K]$

PNOF5	$n_i n_j$	$-\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i = g \vee j = g)$
	$n_i n_j$	$\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i \wedge j \in (F_L, K])$
PNOF6x	$h_i h_j e^{-2S_F}$	$-e^{-S_F}\sqrt{h_i h_j}$	$i \wedge j \in [1, F_L]$
	$\gamma_i\gamma_j/S^x_\gamma$	$-\sqrt{\left[n_i h_j + \frac{\gamma_i\gamma_j}{S^x_\gamma}\right]\left[h_i n_j + \frac{\gamma_i\gamma_j}{S^x_\gamma}\right]}$	$i \in [1, F_L] \wedge j \in (F_L, K]$ $\vee\, j \in [1, F_L] \wedge i \in (F_L, K]$
	$n_i n_j e^{-2S_F}$	$e^{-S_F}\sqrt{n_i n_j}$	$i \wedge j \in (F_L, K]$
PNOF7	$n_i n_j$	$-\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i = g \vee j = g)$
	$n_i n_j$	$\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i \wedge j \in (F_L, K])$
	0	$-\sqrt{n_i h_i n_j h_j}$	$(i \vee j) \in [1, F_L]$ $\wedge [(i \in \Omega_g \wedge j \notin \Omega_g) \vee (j \in \Omega_g \wedge i \notin \Omega_g)]$
	0	$\sqrt{n_i h_i n_j h_j}$	$(i \wedge j) \in (F_L, \infty)$ $\wedge [(i \in \Omega_g \wedge j \notin \Omega_g) \vee (j \in \Omega_g \wedge i \notin \Omega_g)]$

Table 1.2: The expressions for the Δ and Π matrix elements for the PNOF$i, i = 1 \ldots 7$. The diagonal elements for any PNOFi are $\Delta_{ii} = n_i^2$ and $\Pi_{ii} = n_i$. Appearing elements in the table are $h_i = 1 - n_i$, $T_{ij} = n_i n_j - \Delta_{ij}$, and $S_F = \sum_{i=1}^{F_L} h_i$. For PNOF6$x$, $x = d, u, h$ with $S^d_\gamma = \sum_{i=1}^{F_L} \gamma_i$, $S^u_\gamma = \sum_{i>F_L}^{K} \gamma_i$ and $S^h_\gamma = (S^d + S^u)/2$, being $\gamma = n_i h_i + \kappa_i^2 - \kappa_i \sum_{i=j}^{F_L} \kappa_j$ and κ_i taking values of $\kappa_i = h_i e^{-S_F}$ when $i \in [1, F_L]$ and $\kappa_i = n_i e^{-S_F}$ when $i \in (F_L, K]$. For PNOF5, Ω_g is the subspace containing orbital g, which is defined to be below F_L, and several orbitals above F_L.

The benchmark study on which the analysis presented in Chapter 6.3 is based revealed some strengths as well as some problems in the rDMFAs [136]. That study analyzed the ability of the rDMFAs to describe some properties of the harmonium atom, a two-electron physical model. The errors obtained for different properties over increasing correlation regimes revealed an oscillating dependency of the parametrized ML and MLSIC functionals, whereas the rest of functionals showed a linear error increase with the correlation strenght. On the other hand, most of the approximations that yield the correct trace (*i.e.* follow the sum rule) by imposition presented important deviations in the diagonal matrix elements compared with the exact one. This affects the calculation of other chemical properties, for instance, the delocalization index (DI) [168, 169], in which PNOF2, CA, and CGA, for instance, produced errors in predicting DIs above 3%, yet the trace of the 2-DMs are exact.

There is an improvement in PNOFs, where PNOF4 and PNOF6 provided better predictions than older versions of the PNOF considered. Because PNOF4 and PNOF6 reduce to the exact functional for a two-electron system, it is not surprising that they committed the smallest errors compared to the rest of rDMFAs. Whereas most of the functionals performed reasonably well in predicting the electronic repulsion energy, non-negligible errors in the calculation of other properties were obtained.

In Chapter 6.3, the 2-DM approximations presented are built using the exact (within the basis set) NOrbs and NOccs, and thus the orbitals or the occupancies are not optimized for each approximation. Therefore, the errors analyzed in Chapter 5 and discussed in Chapter 6.3 only consider the effect produced by the expression designed to approximate the 2-DM. In a sense, the accuracy of the expressions for the 2-DM approximations is evaluated. This type of error is sometimes called functional-driven error in KS DFT [170], which in that framework evaluates the true error in an exchange-correlation approximation of any given density. Let us introduce the total energy error in rDMFT:

$$\mathrm{E}_{\mathrm{total}} = E\left[\{n_i, \chi_i\}\right] - E^X\left[\left\{n_i^X, \chi_i^X\right\}\right], \tag{1.64}$$

where the first term is the exact energy and the second term is the optimized rDMFA energy. This error can be split in the functional–driven and the 1-rDM–driven error, $\mathrm{E}_{\mathrm{total}} = \mathrm{E}_{\mathrm{F}} + \mathrm{E}_{\mathrm{DM}}$. In the rDMFT framework, the energy functional-driven error is:

$$\mathrm{E}_{\mathrm{F}} = E\left[\{n_i, \chi_i\}\right] - E^X\left[\{n_i, \chi_i\}\right] \tag{1.65}$$

in which the second term in the *r.h.s.* of Eq. 1.65 is the energy of the rDMFA obtained with the exact NOccs and NOrbs – that is, a non-optimized energy. On the other hand, the 1-rDM–driven error (the analogy to the density-driven error in KS DFT [170]) is:

$$\mathrm{E}_{\mathrm{DM}} = E^X\left[\{n_i, \chi_i\}\right] - E^X\left[\left\{n_i^X, \chi_i^X\right\}\right]. \tag{1.66}$$

In this thesis, the 1-rDM–driven (or self-consistent) error is not analyzed.

1.3 The intracule probability density

In chemistry, the concept of an electron pair is fundamental to describe reactivity and binding properties. Electron-pair distribution functions are useful tools that provide relevant chemical information; more precisely, they provide information regarding electron correlation [115]. In the literature, many reviews have been published

regarding the advantages of using distribution functions of 2-PD [123, 171, 172]. SR and LR interactions between two electrons located at $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ can be conveniently analyzed in terms of the center-of-mass vector,

$$\boldsymbol{R} = \frac{1}{2}(\boldsymbol{r_1} + \boldsymbol{r_2}) \tag{1.67}$$

and the interelectronic-distance or relative motion vector

$$\boldsymbol{s} = \boldsymbol{r_1} - \boldsymbol{r_2}. \tag{1.68}$$

Such coordinates define the extracule probability density (EPD)

$$E(\boldsymbol{R}) = \int d\boldsymbol{r_1} d\boldsymbol{r_2} \rho_2(\boldsymbol{r_1}, \boldsymbol{r_2}) \delta\left(\boldsymbol{R} - (\boldsymbol{r_1} + \boldsymbol{r_2})/2\right); \tag{1.69}$$

and the intracule probability density (IPD),

$$I(\boldsymbol{s}) = \int d\boldsymbol{r_1} d\boldsymbol{r_2} \rho_2(\boldsymbol{r_1}, \boldsymbol{r_2}) \delta\left(\boldsymbol{s} - (\boldsymbol{r_1} - \boldsymbol{r_2})\right), \tag{1.70}$$

respectively. Note that the Dirac delta function obliges the evaluation of the function only when $\boldsymbol{s} = \boldsymbol{r_1} - \boldsymbol{r_2}$. The integration over $d\Omega_u$ produces the corresponding radial or isotropic probability density by averaging the probabilities over space [123, 171, 172]:

$$I(s) = s^2 \int d\Omega_s I(\boldsymbol{s}), \tag{1.71}$$

and analogously for the extracule function. $d\Omega_u = \sin\theta_u d\theta_u d\phi_u$ is the solid angle, and u defines either the center-of-mass R or the relative-motion s coordinate. Hereafter, the isotropic probability densities will be simply referred to as probability densities for brevity, assuming that the given probability function is already integrated over the solid angle.

It is assumed that the terms for EPD and IPD were coined by Eddington in its Fundamental Theory series in 1946 [173], yet Debye was the first known scientist to use them as distribution functions of the interelectronic distances [174–176], and were introduced by Coleman in 1967 into the density matrix theory [123]. They provide a genuine advantage with respect to the associated 2-PD, since they are 1-dimensional functions (3-dimensional for the non-averaged distribution function) in contrast to the 6-variable dependency of the spinless 2-PD. Therefore, they permit the analysis of the probability information contained in the 2-PD in a contracted way. Another advantage is that the IPD is the simplest function known in terms of which V_{ee} can be expressed,

$$V_{ee} = \int ds \frac{I(s)}{s}, \tag{1.72}$$

i.e. there is no need to estimate the full 2-PD to retrieve the total energy of a system. In this line, some authors have defined the intracule functional theory [177,178].

Since the EPD is a function of the center-of-mass vector, its evaluation provides information about the location of electron pairs. When evaluated at the origin $R = 0$ it depicts the probability of finding any two electrons on opposite sides with respect to the symmetry center [179–182]. $E(0)$ is named interelectronic counterbalance density, and has been proposed as an indicator of nondynamic correlation in reference [183]. Henceforth, this chapter will focus on the IPD since is an important function for the present thesis. Note that the mathematical properties of the EPD are analogous to the ones of the IPD, which will be explained below. More information about the EPD can be found, for example, in reference [171].

The IPD indicates the probability of two electrons to be separated by a given distance s. The assessment of the IPD at the origin provides information regarding the probability of two electrons to be on top of each other, thus $I(0)$ is known as the electron coalescence density. An interesting fact about the coalescence density is that it can be experimentally determined through integrated total X-ray scattering intensities [184]; in fact, the whole intracule probability density can be obtained from the Fourier-Bessel transform of X-ray scattering cross-sections, being an observable based on the 2-PD [174, 185–189].

The mathematical properties of probability densities are inherited from the 2-PD. They are even functions because the 2-PD is symmetric. The ability to partition the 2-PD into spin contributions (Eq. 1.24) is also acquired,

$$I(s) = I^{\alpha\alpha}(s) + I^{\beta\beta}(s) + I^{\alpha\beta}(s) + I^{\beta\alpha}(s), \tag{1.73}$$

where each intracule component is obtained through the integration of the respective spin component of the 2-PD. The IPD and its spin components receive the normalization factors from the corresponding 2-PD (Eqs. 1.26–1.29) after integration over ds.

Since the IPD is a function of the 2-PD, such 2-PD can be *any* 2-PD, as, for instance, the HF 2-PD (Eq. 1.37). Because of this, we introduce the notation $I^{\mathrm{X}}(s) \equiv I(\rho_2^{\mathrm{X}}, s)$, where X refers to the method to approximate the 2-PD (X = FCI, HF, SD, MBB, ...). This permits an exhaustive analysis of how an approximated pair density correlates electrons [136]. The shape of an IPD is characteristic of each electronic system, yet they always present a particular profile. It is composed

of several maxima centered at the most probable interelectronic distances $\langle s \rangle$, corresponding to the distances between nuclei positions, and followed by a smooth decay to zero at infinite interelectronic separations. Specifically in homonuclear dimers, the profile is characterized by two maxima, one at the SR caused by electron pairs within a nucleus, and another one centered around the interatomic distance, $s \approx R$. As an example, Fig. 1.7 depicts the IPD for the helium dimer.

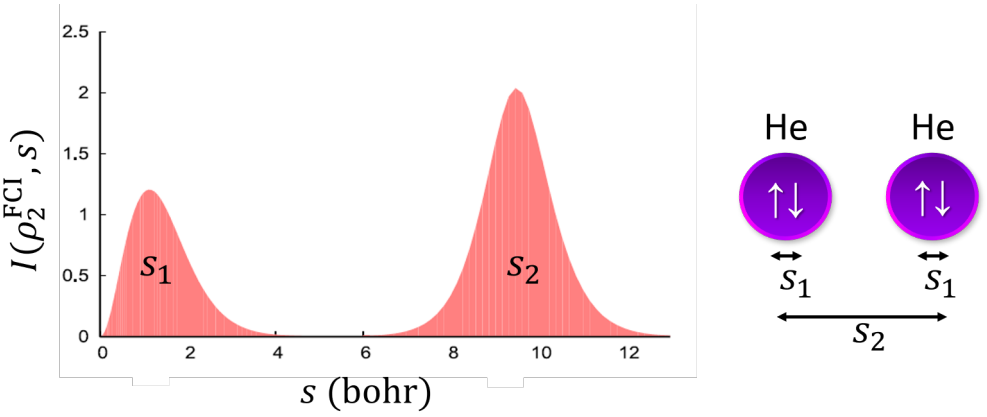


Figure 1.7: Left) The IPD of two helium atoms separated a distance of $R = 9.45$ bohr. Right) A schematic representation of the two different interelectronic distances within the helium dimer. The peak in Left) labelled with s_1 depicts the short-range interelectronic distance for electrons within the same nucleus, and s_2 represents the long-range interelectronic distance for electrons of different nuclei, in which $s_2 \equiv R$.

The usefulness of the presented electron pair-distributions has brought interest amongst the community, since it permits topological insight in the electron distribution [190–192]. $I(0)$ and $E(0)$ have contributed into the understanding of the nature of electron correlation [183, 193]. However, the most widely known application of the IPD may be their use to obtain electron correlation holes [192, 194–202], which are introduced in the following section.

1.3.1 Electron correlation holes

Correlating the motion of electrons implies that the presence of an electron will influence the vicinity of other electrons. The simple existence of an electron implies a drop of probability into finding another electron nearby. In other words, an electron causes a region of decreased probability in its surroundings, a "hole", caused by the Pauli principle and by the fact that it is a charged particle (that is, the repulsive Coulombic and exchange interactions) [25, 27, 171, 203]. There exist two main definitions for correlation holes in quantum chemistry, where the most extended

definition is the one attributed to Coulson, which is an analogy to the definition of electron correlation provided by Löwdin (Eq. 1.1). The second definition is the one proposed by McWeeny, based on statistical probabilities. The definition considered for Coulomb holes in this thesis is the one alluded to Coulson, yet a summary on McWeeny's definition is also given for completeness.

1.3.1.1 Coulson's definition of the Coulomb hole

The definition provided by Coulson is based on the radial intracule probability density. As for the definition of E_c (Eq. 1.1), the correlation hole is defined as the difference between the exact and the HF IPDs [204, 205],

$$h_c(s) = I^{\text{exact}}(s) - I^{\text{HF}}(s) \equiv I(\rho_2^{\text{FCI}}, s) - I(\rho_2^{\text{HF}}, s) \tag{1.74}$$

where, as in the case for the exact nonrelativistic energy in Eq. 1.1, the exact intracule is calculated with the FCI 2-PD for a given basis set. Since the HF 2-PD (Eq. 1.37) includes exchange interactions, the difference in Eq. 1.74 pictures the lack of Coulomb correlation in the HF approach and, therefore, the correlation hole is known as Coulomb hole. The integration of the Coulomb hole is zero, $\int_0^\infty h_c(s)ds = 0$, since both intracule probabilities are normalized with the same coefficient. This property illustrates a physical interpretation of the Coulomb hole, implying that there is a fixed amount of electrons that are shifted from short interelectronic distances to larger ones.

Whereas Debye was the first scientist to advert the Coulomb hole around 1920 [174–176], the first explicit calculation was held in 1961 by Coulson and Neilson [204]. Coulomb holes provide information about changes suffered by the 2-PD when the correlation interactions are turned on. In general, the repulsive interactions cause electrons to increase their expected interelectronic separation $\langle s \rangle$. This phenomenon is portrayed in the IPDs, where the maximum in the exact intracule $I^{\text{exact}}(s)$ is shifted to larger interelectronic distance values s with respect to $I^{\text{HF}}(s)$. The difference between both IPDs results in the well-known shape of Coulomb holes, as portrayed in Fig. 1.8, and it is a physical representation of the improper handling of electrostatic interactions in the HF method as a consequence of the mean-field approximation. One can also consider the Coulomb hole as the intracule of the correlation 2-PD (Eq. 1.30).

The existence of the Coulomb hole implies the existence of a Fermi hole, which reflects the repulsive interactions caused by the fermionic nature of electrons. The

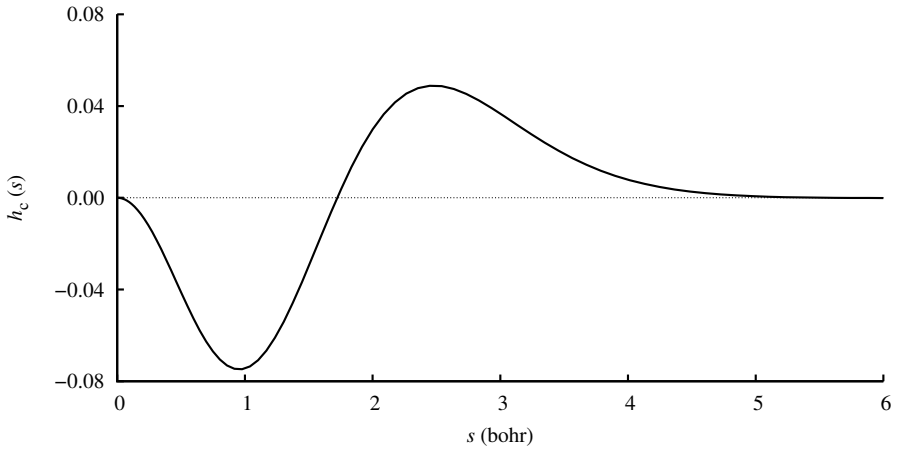


Figure 1.8: Coulomb hole (Eq. 1.74) of the hydrogen molecule H_2 at $R = 1.32$ bohr of bond separation.

Fermi hole, also known as spin correlation hole, can only be defined for electron pairs with like-spin. It represents the decreased probability of two electrons with the same spin to be close to each other. This defines a Fermi hole per each spin, that is, for $\alpha\alpha$ and $\beta\beta$ electron pairs. The definition of Fermi holes was proposed by Boyd and Coulson [206]. The HF method considers exchange interactions but not the Coulombic ones between electrons of opposite spin, being a reference wavefunction that only includes Fermi correlation. The uncorrelated reference must neglect Fermi correlation (a wavefunction that is not antisymmetric). Coulson and Boyd chose the Hartree product wavefunction (Eq. 1.34), yielding to the Fermi hole definition

$$h_{\mathrm{x}}^{\sigma}(s) = I^{\mathrm{HF}}(s) – I^{\mathrm{H}}(s) \equiv I(\rho_2^{\mathrm{HF}}, s) - I(\rho_2^{\mathrm{H}}, s), \tag{1.75}$$

being σ either α or β. Note that the effects of Coulomb correlation are more long-ranged than those of Fermi correlation because the repulsion between a pair of like-spin electrons increases when they approach each other.

1.3.1.2 McWeeny's definition of the Coulomb hole

McWeeny's proposal for electron correlation holes is statistically motivated. This implies that no reference wavefunction is needed, in contrast to Coulson's Coulomb hole [118]. Considering the exchange-correlation density defined in Eq. 1.33,

$$\rho_{\mathrm{xc}}(\mathbf{1}, \mathbf{2}) = \rho(\mathbf{1})\rho(\mathbf{2}) - \rho_2(\mathbf{1}, \mathbf{2}),$$

the exchange-correlation hole or hole density is defined by dividing the exchange-correlation density by the electron density,

$$h_{\mathrm{xc}}(\mathbf{2}|\mathbf{1}) = \frac{\rho_{\mathrm{xc}}(\mathbf{1},\mathbf{2})}{\rho(\mathbf{1})} = \frac{\rho_2(\mathbf{1},\mathbf{2})}{\rho(\mathbf{1})} - \rho(\mathbf{2}). \tag{1.76}$$

The exchange-correlation hole measures how much the probability of finding an electron at $\boldsymbol{r_2}$ changes due to the presence of another one at $\boldsymbol{r_1}$. McWeeny suggested a splitting of the hole density into the exchange hole and the correlation hole, $h_{\mathrm{xc}}(\boldsymbol{r_2}|\boldsymbol{r_1}) = h_{\mathrm{x}}(\boldsymbol{r_2}|\boldsymbol{r_1}) + h_{\mathrm{c}}(\boldsymbol{r_2}|\boldsymbol{r_1})$, according to the spin of the electron pairs. The Fermi hole arises when like-spin cases are considered,

$$h_{\mathrm{x}}(\boldsymbol{r_2}|\boldsymbol{r_1}) = \frac{\rho_2^{\alpha\alpha}(\boldsymbol{r_1},\boldsymbol{r_2})}{\rho^{\alpha}(\boldsymbol{r_1})} - \rho^{\alpha}(\boldsymbol{r_2}) \tag{1.77}$$

and represents the like-spin component of the exchange-correlation hole. Instead, the Coulomb hole is the complete opposite-spin component,

$$h_{\mathrm{c}}(\boldsymbol{r_2}|\boldsymbol{r_1}) = \frac{\rho_2^{\alpha\beta}(\boldsymbol{r_1},\boldsymbol{r_2})}{\rho^{\alpha}(\boldsymbol{r_1})} - \rho^{\beta}(\boldsymbol{r_2}). \tag{1.78}$$

It is usually conceived that the Fermi hole takes into account the effect of the correlated motion of $\alpha\alpha$ electron pairs, and the Coulomb hole describes the correlation between $\alpha\beta$ electron pairs. This, however, is not be completely correct. Certainly, at the HF level, the interactions involving $\alpha\alpha$ electron pairs are the only ones that are correlated. Going beyond HF implies introducing correlation interactions affecting both $\alpha\beta$ and $\alpha\alpha$ spin interactions. Therefore, the repulsion felt between a pair of electrons is increased regardless of their spin. Therefore, the Coulomb hole depicts the correlated motion for any electron pair, for both $\alpha\alpha$ and $\alpha\beta$ pairs. To emphasize this concept, consider Eq. 1.73, where the IPD is composed of both spin contributions. By introducing Eq. 1.73 into Coulson's Coulomb hole definition (Eq. 1.74), it is clear to see that the Coulomb hole depicts the correlation of both $\alpha\alpha$ and $\alpha\beta$ electron pairs. This misconception may be caused by the fact that exchange interactions are stronger than the same-spin electrostatic ones. Most of the $\alpha\alpha$ electron correlation is caused by Fermi correlation, and, generally, the $\alpha\beta$ Coulomb correlation is larger than the $\alpha\alpha$ analogue.

1.4 Separation of dynamic and nondynamic correlation

In section 1.1.1, special attention has been put to emphasize the fact that there is no clear boundary between dynamic and nondynamic correlation [59]. The mixing

of low-lying energetic states with the HF one will always introduce both correlation types into the wavefunction. However, electronic structure methods are in general specialized at improving the description of one of these two types. Many attempts have been done for proposing qualitative correlation expressions to separate dynamic and nondynamic correlation [7,29,183,193,207–217] and, since electron correlation is usually interpreted as an energetic attribute (E_c, Eq. 1.1), most of the partitioning definitions are energy-based. In this section, some relevant partitions of the electron correlation are introduced.

1.4.1 Energy-based separation schemes

The decomposition presented by Mok, Neumann and Handy takes advantage of the ability of the CASSCF method to include nondynamic correlation into the HF reference configuration [29]. They defined nondynamic correlation energy as the fraction of energy that full valence CASSCF is able to describe,

$$E_{\mathrm{NDC}}^{\mathrm{Handy}} = E[\Psi^{\mathrm{CASSCF}}] - E[\Psi^{\mathrm{HF}}] \tag{1.79}$$

and, by completeness, the dynamic correlation energy reads

$$E_{\mathrm{DC}}^{\mathrm{Handy}} = E[\Psi^{\mathrm{exact}}] - E[\Psi^{\mathrm{CASSCF}}]. \tag{1.80}$$

Cioslowski introduced a different approach to provide energetic expressions for correlation components. Cioslowski's idea is to construct a FCI wavefunction that reproduces the HF density by means of a density-constrained approach. This so-called density-constrained FCI energy, $E_{\mathrm{DCFCI}} = E[\Psi^{\mathrm{exact}}, \rho_{\mathrm{HF}}]$, is used to define the dynamic and nondynamic correlation energies [208]:

$$E_{\mathrm{DC}}^{\mathrm{Cios}} = E[\Psi^{\mathrm{exact}}, \rho_{\mathrm{HF}}] - E[\Psi^{\mathrm{HF}}, \rho_{\mathrm{HF}}] \tag{1.81}$$

and

$$E_{\mathrm{NDC}}^{\mathrm{Cios}} = E[\Psi^{\mathrm{exact}}, \rho_{\mathrm{exact}}] - E[\Psi^{\mathrm{exact}}, \rho_{\mathrm{HF}}], \tag{1.82}$$

respectively. Considering a wavefunction where the HF reference state does not strongly mix with other configurations, the HF electron density is expected to bring a correct approximation to the exact density. According to this statement, the nondynamic energy should then be close to zero [208].

On the other hand, Valderrama, Ludeña, and Hinze proposed a generalization of Cioslowski's scheme by presenting the former decomposition divided into two paths [209, 210]. One of the paths, which the authors call path I, corresponds to

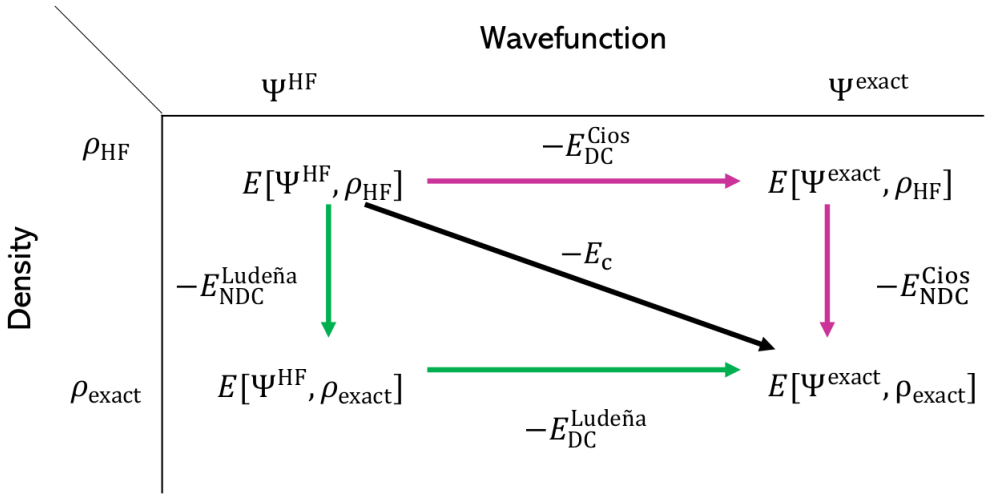


Figure 1.9: Schematic diagram of E_c decomposition into dynamic and nondynamic correlation components, according to Cioslowski's (purple path, Eqs. 1.81 and 1.82), and Ludeña and coworkers (green path, Eqs. 1.83 and 1.84) proposals. Adapted from Ref. [209].

the decomposition proposed by Cioslowski; path II, instead, introduces a different intermediate energy component, which corresponds to a HF calculation using a single-determinant wavefunction that is restricted to reproduce the exact density, namely $E_{HF}[\rho_{exact}]$. Therefore, the definition proposed for the dynamic correlation energy reads

$$E_{DC}^{Ludeña} = E[\Psi^{exact}, \rho_{exact}] - E[\Psi^{HF}, \rho_{exact}], \tag{1.83}$$

and, for the nondynamic counterpart,

$$E_{NDC}^{Ludeña} = E[\Psi^{HF}, \rho_{exact}] - E[\Psi^{HF}, \rho_{HF}]. \tag{1.84}$$

Fig. 1.9 contains a representation of the paths to retrieve the E_c, according to the definitions proposed by Cioslowski (path I) and Ludeña and coworkers (path II). $-E_{DC}^{Cios}$ and $-E_{NDC}^{Cios}$ (upper triangular elements in Fig. 1.9) are positive quantities, since each step in path I implies a relaxation of the system that leads to lower energies. On the other hand, the first step in path II is destabilizing ($-E_{NDC}^{Ludeña} < 0$) because the HF wavefunction does not bear the HF optimal density anymore. Instead, describing the system with the exact wavefunction makes $-E_{DC}^{Ludeña}$ positive, and is the stabilizing step of path II.

Path II is inspired by the definitions proposed by Gross, Petersilka, and Grabo [10], and the ones by Görling and Ernzerhof, all of them in the framework of KS DFT [218]. Valderrama *et al.* defined the nondynamic correlation energy in path II as the sum of the concepts presented by the former authors. Their works attempted to provide an actual comparison between electron correlation energies in

KS DFT and wavefunction theory. In one hand, Görling and Ernzerhof compared the energy arisen from the difference between the KS and the HF wavefunctions yielding the correct electron density, $\Delta E = E[\Psi^{\mathrm{HF}}, \rho_{\mathrm{exact}}] - E[\Psi^{\mathrm{KS}}, \rho_{\mathrm{exact}}]$ [218]. On the other side, Gross, Petersilka and Grabo discussed the different definitions for E_{c} in wavefunction theory (which they name "quantum chemistry correlation energy") and the one in the KS DFT framework. Whereas the uncorrelated energy for the wavefunction theory definition is the same proposed by Löwdin, being the HF solution obtained with the self-consistent HF orbitals, the uncorrelated energy in the DFT framework is the exact KS solution. The difference between both electron correlation definitions results in $\Delta = E[\Psi^{\mathrm{KS}}, \rho_{\mathrm{exact}}] - E[\Psi^{\mathrm{HF}}, \rho_{\mathrm{HF}}]$ [10].

1.4.2 Electron correlation separation based on natural occupancies

In 2016, Ramos-Cordoba, Salvador and Matito presented an electron correlation decomposition scheme in which dynamic and nondynamic correlation are separated using solely NOccs [219]. This approach may be more convenient than an energetic decomposition of electron correlation since it can be used in theoretical approaches and obtain the correlation separation of other observables.

It has been mentioned throughout this thesis that the 2-PD is a quantity that should contain all the electron correlation information of the system [115]. Eqs. 1.40 and 1.41 introduced in section 1.2.2 show that the HF 2-DM can be expressed in terms of occupation numbers and, when NOccs from a correlated calculation are used instead of the HF ones, the resulting 2-DM is the SD approximation to the 2-DM (Eq. 1.42). To use the same labeling appearing in the section 1.4.1, the SD and HF 2-DM are referred as ${}^2\mathrm{D}^{\mathrm{SD}} \equiv {}^2\mathrm{D}^{\mathrm{SD}}[{}^1\mathrm{D}]$ and ${}^2\mathrm{D}^{\mathrm{HF}} \equiv {}^2\mathrm{D}^{\mathrm{SD}}[{}^1\mathrm{D}^{\mathrm{HF}}]$, where the square brackets indicate that the 2-DM is a functional of the 1-DM.

Inspired by the energy-based decompositions previously introduced, more precisely by the second path presented by Valderrama, Ludeña and Handy, the authors suggested a correlation decomposition in terms of the 2-PD in the NOrb basis. In their work, they decompose the correlation 2-PD ($\boldsymbol{C}_{ij}^{kl}$) into two matrices [219],

$$\boldsymbol{C}_{ij}^{kl} = \Lambda_{ij}^{kl} + \Gamma_{ij}^{kl}, \tag{1.85}$$

being

$$\Lambda_{ij}^{kl} = {}^2\mathrm{D}_{ij}^{kl,\mathrm{SD}} - {}^2\mathrm{D}_{ij}^{kl,\mathrm{HF}} = (n_i n_j - n_i^{\mathrm{HF}} n_j^{\mathrm{HF}})(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}); \tag{1.86}$$

$$\Gamma_{ij}^{kl} = {}^2\mathrm{D}_{ij}^{kl} - {}^2\mathrm{D}_{ij}^{kl,\mathrm{SD}} = {}^2\mathrm{D}_{ij}^{kl} - n_i n_j(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}). \tag{1.87}$$

Λ_{ij}^{kl} is the difference between the SD and HF 2-PD and it reduces to a difference of correlated and uncorrelated NOccs, since they share the same expression but use a different ingredient. Since multireference systems (dominated by nondynamic correlation) present large deviations of 0 and 1 in their NOccs, Λ_{ij}^{kl} will be significantly large in such cases and practically zero when the system has a monodeterminantal picture. Γ_{ij}^{kl} depicts the difference between the exact and the SD approximation to the pair density. Both 2-PD involve the same NOccs, but the expression to depict the 2-PD is different. This definition is coincident with the one for the cumulant matrix of the 2-DM, presented in Eq. 1.61 [124, 125].

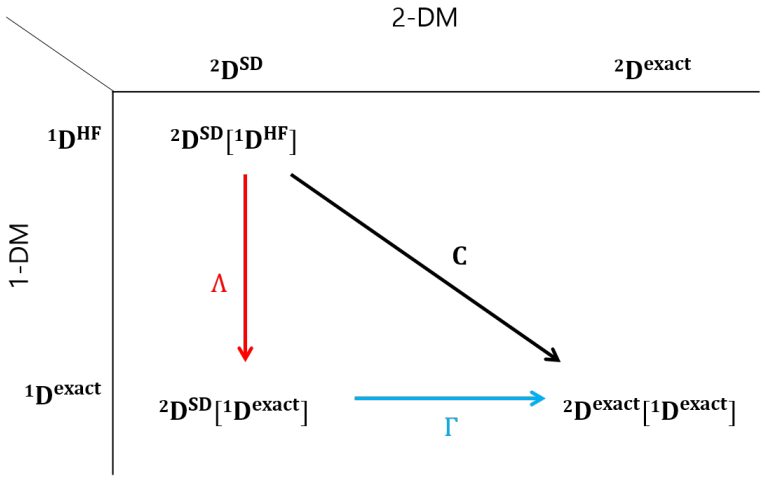


Figure 1.10: Schematic diagram of the correlation 2-DM C_{ij}^{kl} decomposition to Λ_{ij}^{kl} and Γ_{ij}^{kl} components to account for dynamic and nondynamic correlation, in terms of the 1-DM. Adapted from Ref. [219].

The hypothesis of this framework is that large changes in NOccs are caused by nondynamic correlation being present in a system, caused by the configurations that mix strongly with the HF determinant in the CI wavefunction. It is then expected that Λ_{ij}^{kl} and Γ_{ij}^{kl} should account for nondynamic and dynamic correlation, respectively. From a minimal-basis model, the authors proposed a set of correlation indicators based on Eqs. 1.86 and 1.87 [219],

$$I_{\mathrm{D}} = \frac{1}{4}\sum_{\sigma,i}\sqrt{n_i^\sigma\,(1-n_i^\sigma)} - \frac{1}{2}\sum_{\sigma,i} n_i^\sigma\,(1-n_i^\sigma)\,, \tag{1.88}$$

$$I_{\mathrm{ND}} = \frac{1}{2} \sum_{\sigma,i} n_i^\sigma \left(1 - n_i^\sigma\right), \tag{1.89}$$

and

$$I_{\mathrm{T}} = I_{\mathrm{D}} + I_{\mathrm{ND}} = \frac{1}{4} \sum_{\sigma,j} \sqrt{n_i^\sigma \left(1 - n_i^\sigma\right)}, \tag{1.90}$$

which provide simple measures of both correlation components and total correlation in terms of NOccs and permit the evaluation of correlation for any type of calculation. Note that I_{ND} corresponds to the deviation from idempotency of the 1-rDM [220]. Fig. 1.11 contains an adaptation of a 2-electron 2-orbital model that shows that I_{D} is larger than the nondynamic correlation indicator in extremelly small NOcc values, but I_{ND} takes over as the NOcc increases. The maximum value of I_{ND} is obtained when one electron is split between two orbitals, which gives a maximum of nondynamic correlation.

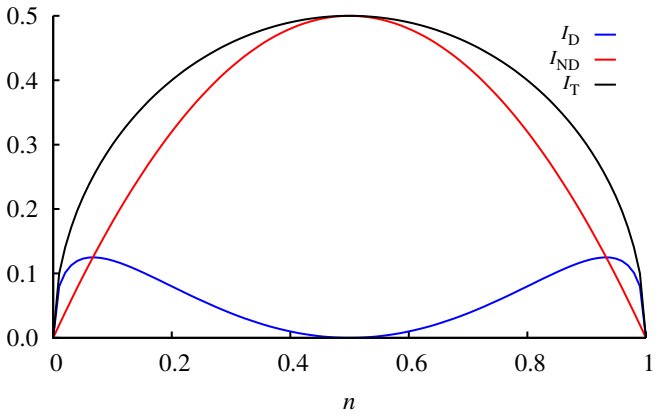


Figure 1.11: The dynamic I_{D} (blue), nondynamic I_{ND} (red) and total I_{T} (black) global electron correlation indicators (Eqs. 1.88–1.90) for a homonuclear minimal-basis two-electron model in singlet state, for different NOccs n. Adapted from Ref. [219].

These global indicators [219] were adapted to local indicators of electron correlation [221]. By assuming a homogeneous contribution of the orbitals, these indices provide measures of electron correlation in particular points in space. The expressions for the global indicators (Eqs. 1.88–1.90) can be transformed into local expressions by multiplication of the spin-natural orbitals [221]:

$$I_{\mathrm{D}}(\boldsymbol{r}) = \frac{1}{4} \sum_{\sigma,i} \sqrt{n_i^\sigma \left(1 - n_i^\sigma\right)} - \frac{1}{2} \sum_{\sigma,i} n_i^\sigma \left(1 - n_i^\sigma\right) |\chi_i^\sigma(\boldsymbol{r})|^2, \tag{1.91}$$

$$I_{\mathrm{ND}}(\boldsymbol{r}) = \frac{1}{2} \sum_{\sigma,i} n_i^\sigma \left(1 - n_i^\sigma\right) |\chi_i^\sigma(\boldsymbol{r})|^2, \tag{1.92}$$

$$I_{\mathrm{T}}(\boldsymbol{r}) = \frac{1}{4}\sum_{\sigma,j}\sqrt{n_i^\sigma\,(1-n_i^\sigma)}|\chi_i^\sigma(\boldsymbol{r})|^2. \tag{1.93}$$

The integration of the local indicators over the space coordinates results into the corresponding global indicator, $\int I_{\mathrm{X}}(\boldsymbol{r})d\boldsymbol{r} = I_{\mathrm{X}}$, being I_{X} any of the three indicators.

Chapter 2

Objectives

This thesis is an extension to the work presented by Ramos-Cordoba, Salvador, and Matito [219, 221], introduced in section 1.4.2, and also the rDMFT benchmarking study of reference [136].

The motivation to continue working on the correlation decomposition relies on the promising predictions obtained with both global and local correlation indicators [219, 221]. On the other hand, there is a no comprehensive benchmarking of rDMFT approximations in the literature that considers a qualitatively large set of molecules and tests. The objectives of this thesis are separated into two principal points, 1) the study of the correlated part of the 2-PD, its separation into two components, and posterior analysis by range utilizing Coulomb holes, and 2) the study and benchmark of approximate 2-DM within the rDMFT framework. Note that the objectives of this thesis are connected, as the quantity that is divided into two components is the same quantity that is approximated in rDMFT.

The first objective concerns the study of the correlated part of the 2-PD, $\Delta\rho_2^c$, and the proposal of a correlation decomposition scheme that separates $\Delta\rho_2^c$ into two components. This would be the first step to fulfill the first objective. The decomposition to be presented in this thesis is based on the correlation decomposition introduced in section 1.4.2. Such decomposition was simplified in order to reach an expression in terms of NOccs and NOrbs [219,221]. Instead, in this thesis no simplification is made in order to provide an exact decomposition. This partition pretends to analyze a property (the 2-PD) rather than a product quantity (the electronic energy). Whereas the energy is the direct output from a quantum chemical calculation, the 2-PD is the natural ingredient to obtain the electron-electron repulsion. The goal of the decomposition is to gather enough information of the correlation part of the 2-PD and its correlation components in order to reproduce their behavior in

physical models. The long-term goal of this project (of which this thesis represents the first step) is to produce new exchange-correlation functionals for KS DFT. The 2-PD components, named $\Delta\rho_2^{c_I}$ and $\Delta\rho_2^{c_{II}}$, are responsible for capturing dynamic and nondyamic correlation in some extent, and mainly contain either long- or short-range interactions, respectively. They are introduced and analyzed in Chapter 3, in two articles, and discussed in Chapter 6.1.

The second step of the first objective of this thesis is to validate the 2-PD decomposition by range-separating electron correlation. Coulomb holes are useful tools to do this analysis since they provide a clear vision of the electron correlation effects by range. Because intracule pair densities (IPDs) (from which Coulomb holes are defined) are functionals of the 2-PD, the correlation decomposition can be applied to Coulomb holes. This permits an analysis of $\Delta\rho_2^{c_{II}}$ and $\Delta\rho_2^{c_I}$ also by range. Dynamic correlation is mainly short-ranged in nature, whereas nondynamic correlation is most important when electrons are largely separated (long-ranged), arising from entanglement and orbital degeneracies emerging from molecular dissociations. Nevertheless, nondynamic correlation can also stem from short-range interactions as well generated by orbital degeneracies (consider the beryllium atom as an example). Dynamic correlation can also induce important effects in long-range interactions due to London dispersion forces. Therefore, this step pretends to validate whether the c_I and c_{II} correlation components are capable of describing these short- and long-range correlation effects, and determine their ability to account for nondynamic and dynamic correlation. The analysis is carried out through molecular systems and physical models. This point is treated together with the first one in the same chapters of this thesis.

The third and last point to fulfill the fist objective involves the analysis of London dispersion interactions. Because these interactions arise from long-range dynamic correlation, the long-range region of the correlation component responsible for capturing dynamic correlation effects is evaluated. Special attention is put into characterizing dispersion interactions in the IPD. An analytic study in the perturbation theory framework is carried out in order to obtain a universal signature for dispersion interactions, which is reminiscent of the well-known dispersion energy decay with the interfragment distance R, $E^{\text{disp}} \propto R^{-6}$. This property of dispersion interactions is also examined in the range-separation of the Coulomb hole. The motivation for this point is to find a property (other than the energy) that can provide a connection with dispersion interactions. Because the IPD is a functional of the 2-PD, the dispersion condition can be linked with the 2-PD. Obtaining a dispersion

condition for the 2-PD would provide with the ability to design new models that could describe the correct decay of dispersion interactions with R. The assessment of this objective is addressed in Chapter 4 and examined in Chapters 6.2 and 6.1.2.1.

The second main objective of this thesis is to provide a benchmark for a group of approximations of the reduced density matrix functional theory (rDMFT) by comparison with FCI references. The tests used to analyze the accuracy of the functionals involve energetic predictions, chemical properties, and known conditions that the structure of a 2-DM must fulfill. Some benchmark studies for rDMFT approximations are already present in the literature, but most of the tests performed in those works are based on energy predictions, and they do not consider a large list of rDMFAs. The set of molecules in which the approximations are applied is also small and, hence, we believe that this far there is no exhaustive benchmark for rDMFT approximations in the literature. Chapter 5 in this thesis pretends to complement these former studies through a set of 14 diatomic molecules and other systems to measure the accuracy of 17 rDMFT approximations via tests that go beyond energy benchmarks. Results are discussed in Chapter 6.3.

To sum up, the objectives of this thesis are:

1) To study of the correlated part of the pair density (2-PD).

   1.1 To separate the correlated part of the 2-PD into two correlation components.

   1.2 To study the range-separation of the correlated part of the 2-PD and the decomposed correlation components using intracule probability densities (IPDs) and Coulomb holes. Validate the correlation decomposition.

   1.3 To study dispersion interactions using the range-separation of the correlation decomposition. Connect the IPD with dispersion interactions.

2) To provide an exhaustive rDMFT benchmark with molecules, evaluating the predictive ability of chemical properties besides the energy, and analyzing the expressions designed for the functionals.

Chapter 3

The pair density-based correlation decomposition of the Coulomb hole

3.1 The separation of the c_I and c_{II} correlation components in the Coulomb hole, I

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

Cite This: J. Phys. Chem. Lett. 2019, 10, 4032−4037

pubs.acs.org/JPCL

Singling Out Dynamic and Nondynamic Correlation

Mireia Via-Nadal,[†,‡,∥] Mauricio Rodríguez-Mayorga,[†,‡,¶,∥] Eloy Ramos-Cordoba,[†,‡] and Eduard Matito*[,†,§]

[†]Donostia International Physics Center (DIPC), 20018 Donostia, Euskadi, Spain

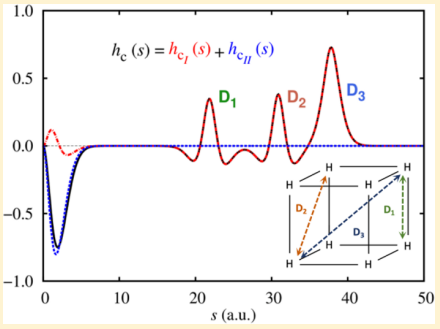
[‡]Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Euskadi, Spain

[¶]Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, C/Maria Aurèlia Capmany, 69, 17003 Girona, Catalonia, Spain

[§]IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Euskadi, Spain

Supporting Information

ABSTRACT: The correlation part of the pair density is separated into two components, one of them being predominant at short electronic ranges and the other at long ranges. The analysis of the intracular part of these components permits to classify molecular systems according to the prevailing correlation: dynamic or nondynamic. The study of the long-range asymptotics reveals the key component of the pair density that is responsible for the description of London dispersion forces and a universal decay with the interelectronic distance. The natural range-separation, the identification of the dispersion forces, and the kind of predominant correlation type that arise from this analysis are expected to be important assets in the development of new electronic structure methods in wave function, density, and reduced density-matrix functional theories.



Electron correlation being the Holy Grail of electronic structure methods, it has been the subject of extended analysis.[1−14] The solution of quantum many-body problems hinges on the type of correlation present in the system, and one of the most practical classifications consists in the separation between dynamic- and nondynamic-correlation-including methods. Indeed, there are accurate methods to study systems with one predominant correlation type, but systems presenting both correlation types pose one of the greatest current challenges in electronic structure theory.[15−17]

The attempt at taking the best of both worlds has led to a resurgence of interest in hybrid schemes,[18] merging methods that recover different correlation types.[19−21] Among hybrid implementations, the most successful one is based on the range separation of electron correlation,[18,22,23] using a mixing function to combine approximations that account for short-range dynamic correlation—such as density functional approximations—with approaches providing correct long-range asymptotics. The performance of these methods pivots on the choice of the function combining the two approaches, which provides a natural splitting of the Coulomb interaction and thus the pair density.[24] In range-separation approximations, the typical choice is the error function that, in turn, depends on an attenuating parameter, which is both system- and property-dependent.[25,26] Even though the methods are chosen according to their ability of recovering dynamic and nondynamic correlation, the range separation of the pair density has not been motivated by the correlation type present in the system, risking double counting of electron correlation.

Thus far, there has been very few attempts to separate dynamic and nondynamic correlation,[2,4−9,13,27] most of them based on energy calculations. The lack of a physically sound separation of dynamic and nondynamic correlation precludes individual treatment of these effects. We analyze the decomposition of the pair density into three components: the uncorrelated reference and two correlation terms. The latter two behave differently with respect to large changes of the first-order reduced density matrix (1-RDM), permitting the identification of systems with prevalent dynamic or nondynamic correlation.[2,6,7,11,12] Some of us have recently used a similar strategy to obtain scalar[11] and local[12] measures of dynamic and nondynamic electron correlation from a two-electron model. The intracule of the correlation components of the pair density yields a 2-fold separation of the Coulomb hole in terms of correlation type and interelectronic range. These components of the pair density display a simpler mathematical form than the total pair density, one of them being dominant at short ranges and one with prevailing long-range contributions. This feature is particularly convenient for the design of energy functionals in wave function, density, and density matrix functional theories. As a result of this separation, we will clearly identify the part of the pair density that is responsible for the correct description of van der Waals interactions and unveil a universal condition it should satisfy.[28] To our knowledge, the

Received: May 14, 2019
Accepted: June 17, 2019
Published: June 17, 2019

Downloaded via UNIV DEL PAIS VASCO on November 19, 2019 at 10:40:36 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

© 2019 American Chemical Society

DOI: 10.1021/acs.jpclett.9b01376
J. Phys. Chem. Lett. 2019, 10, 4032−4037

latter is the only known condition of the pair density that can be employed to design methods including van der Waals interactions.

Theoretical Background. Let us consider the pair density of an N-electron system described by the $\Psi(\mathbf{1}, ..., \mathbf{n})$ wave function

$$\rho_2(\mathbf{1}, \mathbf{2}) = \frac{N(N-1)}{2} \int d_3 ... d_n |\Psi(\mathbf{1}, ..., \mathbf{n})|^2 \quad (1)$$

where numerical variables $(\mathbf{1}, \mathbf{2}, ...)$ refer to space and spin coordinates. Upon integration over its coordinates, the pair density can be reduced to the intracule density, which only depends on the interelectronic range separation, s

$$I(\rho_2, s) = \int d_1 d_2 \, \rho_2(\mathbf{1}, \mathbf{2}) \delta(s - r_{12}) \quad (2)$$

where r_{12} is the Euclidean distance between the electrons at $\mathbf{1}$ and $\mathbf{2}$. The intracule density is the simplest function in terms of which we can express the Coulomb interaction energy

$$V_{ee}[I] = \int ds \frac{I(\rho_2, s)}{s} \quad (3)$$

The electron correlation contents of the pair density can be determined by the difference between the actual pair density and an uncorrelated reference, which here we choose to be the Hartree−Fock (HF) one

$$\Delta\rho_2^c(\mathbf{1}, \mathbf{2}) = \rho_2(\mathbf{1}, \mathbf{2}) - \rho_2^{HF}(\mathbf{1}, \mathbf{2}) \quad (4)$$

The intracule of this function is Coulson's Coulomb hole[29]

$$h_c(s) = I(\Delta\rho_2^c, s) = \int d_1 d_2 \Delta\rho_2^c(\mathbf{1}, \mathbf{2}) \delta(s - r_{12}) \quad (5)$$

In order to split the correlation part of the pair density (eq 4), we employ an approximate pair density, the single-determinant (SD) ansatz of the pair density[1]

$$\rho_2^{SD}(\rho_1, \mathbf{1}, \mathbf{2}) = \rho_1(\mathbf{1})\rho_1(\mathbf{2}) - |\rho_1(\mathbf{1}; \mathbf{2})|^2 \quad (6)$$

where $\rho_1(\mathbf{1}; \mathbf{2})$ is the 1-RDM and $\rho_1(\mathbf{1}) \equiv \rho_1(\mathbf{1}; \mathbf{1})$ is the electron density. Substituting ρ_1 by the HF 1-RDM in eq 6 yields the HF pair density, i.e.,

$$\rho_2^{HF}(\mathbf{1}, \mathbf{2}) = \rho_2^{SD}(\rho_1^{HF}, \mathbf{1}, \mathbf{2}) \quad (7)$$

which does not account for electron correlation. However, $\rho_2^{SD}(\rho_1, \mathbf{1}, \mathbf{2})$ can be regarded as an approximation to the actual pair density; an approximation that does not account for dynamic correlation either at short-[30] or at long-range.[28] Figure 1 depicts the two paths of arriving at the exact $\rho_2(\rho_1, \mathbf{1}, \mathbf{2})$ from $\rho_2^{SD}(\rho_1^{HF}, \mathbf{1}, \mathbf{2})$, either straightforwardly or through the intermediate SD approximation. The latter path defines the decomposition of the correlation part of the pair density

$$\begin{aligned}\Delta\rho_2^c(\mathbf{1}, \mathbf{2}) &= (\rho_2^{SD}(\rho_1, \mathbf{1}, \mathbf{2}) - \rho_2^{SD}(\rho_1^{HF}, \mathbf{1}, \mathbf{2})) \\ &\quad + (\rho_2(\rho_1, \mathbf{1}, \mathbf{2}) - \rho_2^{SD}(\rho_1, \mathbf{1}, \mathbf{2})) \\ &= \Delta\rho_2^{c_I}(\mathbf{1}, \mathbf{2}) + \Delta\rho_2^{c_{II}}(\mathbf{1}, \mathbf{2})\end{aligned} \quad (8)$$

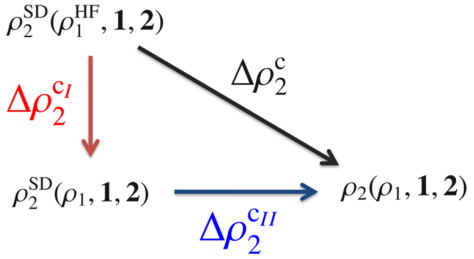


Figure 1. Correlation part of the pair density, $\Delta\rho_2^c$, decomposed into two components.

$\Delta\rho_2^{c_I}(\mathbf{1}, \mathbf{2})$ will be large only if the HF 1-RDM and the actual 1-RDM are significantly different, and in such case, the system will be affected by nondynamic correlation. Indeed, the wave function of systems dominated by dynamic correlation can be described by a large expansion of Slater determinants with one of them (the HF one) having an expansion coefficient very close to one.[31] Therefore, these systems are characterized by a 1-RDM that retains the shape of the HF 1-RDM. Conversely, the wave function of nondynamic-correlated systems can be written as a shorter expansion of Slater determinants, but in this case, the HF determinant has an expansion coefficient that is qualitatively smaller than one.[31] Since the 1-RDM is determined by the square of the expansion coefficients, we expect systems affected by nondynamic correlation to display large $\Delta\rho_2^{c_I}(\mathbf{1}, \mathbf{2})$. Some authors have used similar arguments to use the electron density (the diagonal part of the 1-RDM) as a means to define dynamic and nondynamic correlation energy.[2,6] In this work, we prefer to employ the 1-RDM because the cases of spin entanglement would not be regarded as nondynamic correlation if only density differences were considered. Indeed, in the stretched H_2 molecule, the HF electron density is qualitatively similar to the exact one, whereas there are large and notorious differences between the exact and the HF 1-RDMs.

The magnitude of $\Delta\rho_2^{c_I}(\mathbf{1}, \mathbf{2})$ can be thus regarded as a measure of nondynamic correlation, but it can also be interpreted as the correlation retrieved by using the actual 1-RDM rather than the HF one to construct the pair density. Conversely, $\Delta\rho_2^{c_{II}}(\mathbf{1}, \mathbf{2})$ does not depend on the differences between ρ_1^{HF} and ρ_1, but on the validity of the SD approximation. Note that $\Delta\rho_2^{c_{II}}$ coincides with the cumulant of the pair density.[14,32] The intracule functions of $\rho_2^{SD}(\rho_1, \mathbf{1}, \mathbf{2})$ and the exact pair density, $\rho_2(\rho_1, \mathbf{1}, \mathbf{2})$, display the same asymptotic behavior[33] and, therefore, $\Delta\rho_2^{c_{II}}$ is dominated by the short-range component. Interestingly, $\Delta\rho_2^{c_I}(\mathbf{1}, \mathbf{2})$ is the long-range-dominant component of the correlated part of the pair density (eq 4) because the HF and the exact 1-RDM can differ substantially at large separations, for instance, in the presence of entanglement. On the contrary, $\Delta\rho_2^{c_I}(\mathbf{1}, \mathbf{2})$ displays very small values at small interelectronic distances mostly due to the opposite-spin part of this term.

The current partition

$$\rho_2(\mathbf{1}, \mathbf{2}) = \rho_2^{HF}(\mathbf{1}, \mathbf{2}) + \Delta\rho_2^{c_I}(\mathbf{1}, \mathbf{2}) + \Delta\rho_2^{c_{II}}(\mathbf{1}, \mathbf{2}) \quad (9)$$

provides a natural range separation of the pair density that can be employed to split the Coulomb hole into two correlation components

$$h_c(s) = h_{c_I}(s) + h_{c_{II}}(s) = I(\Delta\rho_2^{c_I}, s) + I(\Delta\rho_2^{c_{II}}, s) \quad (10)$$

naturally yielding a separation of electron correlation by range. We will show that the decay of $I(\Delta\rho_2^{c_{II}}, R)$ is universal and that it corresponds to a characteristic signature of London dispersion forces (R being the distance between two atoms in the molecule).

DOI: 10.1021/acs.jpclett.9b01376
J. Phys. Chem. Lett. 2019, 10, 4032−4037

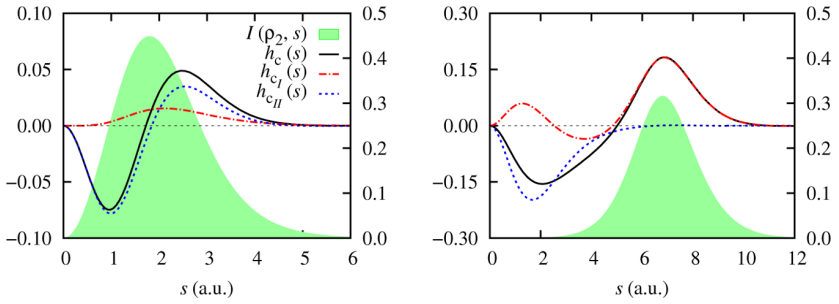


Figure 2. Total Coulomb hole (black), $h_{c_I}(s)$ (red), and $h_{c_{II}}(s)$ (blue) correlation components, and the intracule density (shadowed green region, right y-axis) of the H_2 molecule at 1.32 and 7.56 au bond lengths.

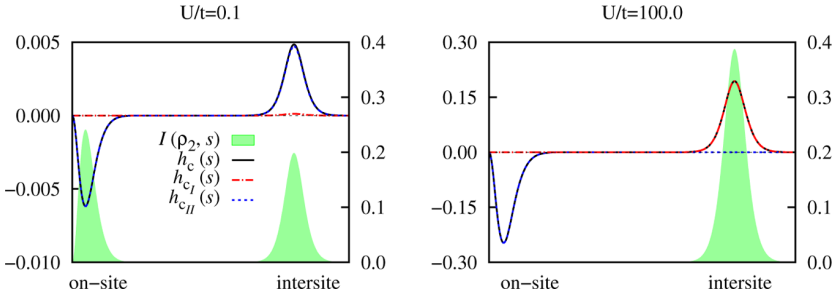


Figure 3. Coulomb hole (see Figure 2 for further details) of the two-site real-space Hubbard dimer for various U/t values.

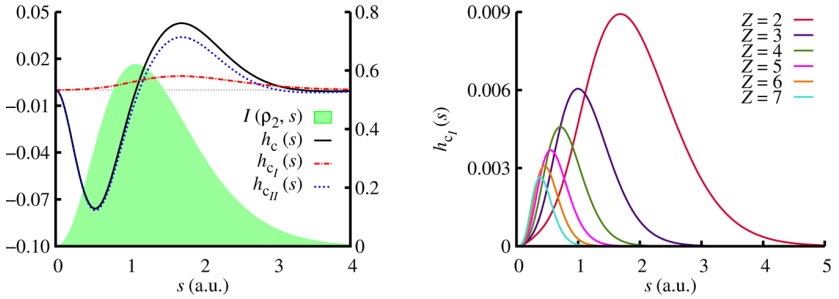


Figure 4. Coulomb hole of He (l.h.s.) and $h_{c_I}(s)$ for the isoelectronic series of He (r.h.s.).

Results and Discussion. In the following, we introduce five selected examples that illustrate the effectiveness of the current scheme to separate the correlation part of the Coulomb hole at different ranges and how the long-range of $\Delta\rho_2^{c_{II}}$ can be used to identify and characterize van der Waals interactions.

Hydrogen Molecule.[34] At the equilibrium geometry, $h_{c_{II}}(s)$ dominates over $h_{c_I}(s)$ at all interelectronic distances s, as shown in the left panel of Figure 2, whereas $h_{c_I}(s)$ increases importantly as the bond is stretched, in line with the expected increase of nondynamic correlation. The most likely distribution of the electron pair at large bond lengths corresponds to one electron sitting at each atom, and accordingly, the intracule density peaks around the bond-length distance. At the dissociation limit, the long-range part of the Coulomb hole is completely determined by $h_{c_I}(s)$ because one isolated electron cannot give rise to dynamic correlation. Hence, the unrestricted HF calculation of H_2 produces Coulomb hole components that are not distinguishable from FCI.[35] A simple interpretation is also obtained from valence bond theory: at large separations, the exact pair density is entirely described by covalent components, whereas the HF pair density contains equally contributing ionic and covalent terms. $h_{c_{II}}(s)$ removes the ionic contribution (i.e., removes contributions keeping the electrons at short distances), whereas $h_{c_I}(s)$ adds the missing covalent contribution (i.e., adds contributions placing one electron in each atom); in accord with the results plotted in the r.h.s. of Figure 2 (see also the Supporting Information).

Hubbard Dimer. The Hubbard dimer is the simplest model of interacting particles in a lattice and conceivably the most studied model for testing methods at different correlation regimes.[36,37] We employ the one-dimension Hamiltonian of the Hubbard model

$$\hat{H} = -t \sum_{\langle\mu,\nu\rangle,\sigma} (\hat{c}^{\dagger}_{\mu\sigma}\hat{c}_{\nu\sigma} + \hat{c}^{\dagger}_{\nu\sigma}\hat{c}_{\mu\sigma}) + U \sum_{\mu} \hat{\rho}_{\mu\alpha}\hat{\rho}_{\mu\beta} \tag{11}$$

where μ and ν denote the sites, σ is the spin polarization (α or β), $\hat{c}^{\dagger}_{\mu\sigma}$ and $\hat{c}_{\mu\sigma}$ are creation and annihilation operators of one electron with spin σ in site μ, and $\hat{\rho}_{\mu\sigma}$ stands for a one-particle number operator with spin σ acting on site μ. t is the hopping parameter, and U is the on-site interaction parameter. These

DOI: 10.1021/acs.jpclett.9b01376
J. Phys. Chem. Lett. 2019, 10, 4032−4037

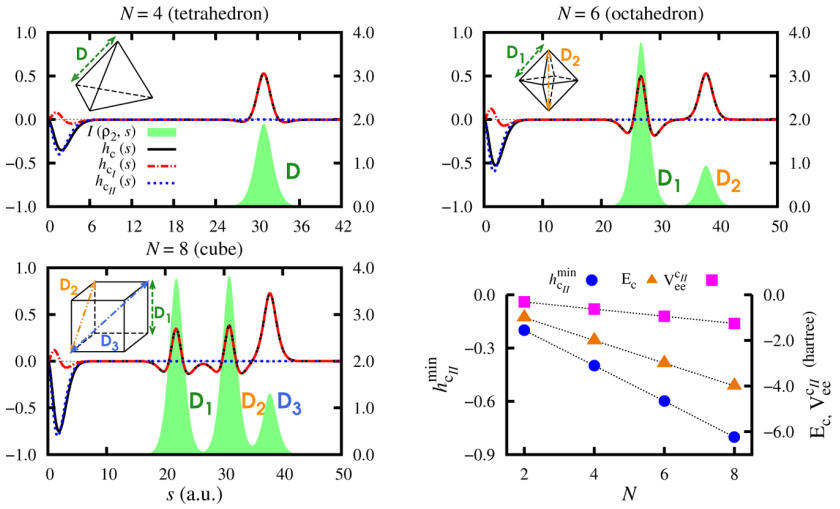


Figure 5. Coulomb hole of N hydrogen atoms. D, D_1, D_2, and D_3 indicate the different distances between the H atoms placed at the vertices of the respective polyhedra. The bottom r.h.s. plot displays the minimal value of $h_{c_{II}}(s)$, $h_{c_{II}}^{min}$, the part of electron−electron repulsion that corresponds to $\Delta\rho_2^{c_{II}}$ (i.e., $V_{ee}^{c_{II}} = \int ds\, h_{c_{II}}(s)/s$), and the correlation energy (E_c), as a function of N.

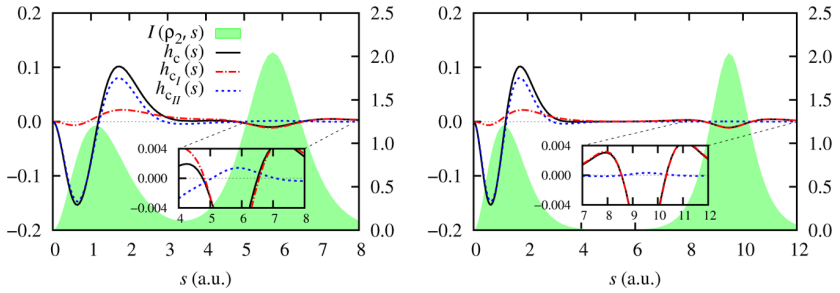


Figure 6. Coulomb hole of the He_2 molecule at two bond lengths (5.67 and 9.45 au, left and right). The inset plots reproduce the ones above on a narrower interval.

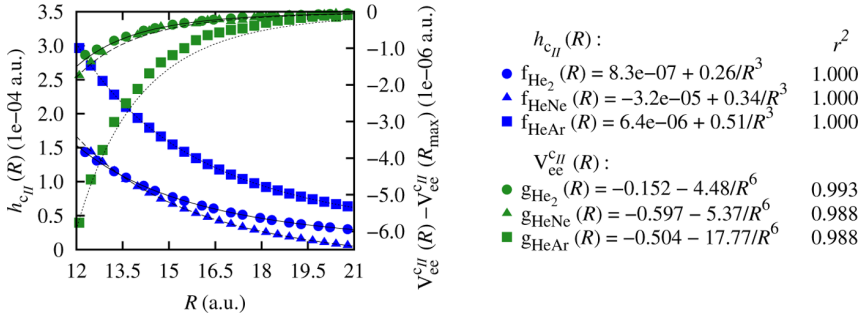


Figure 7. $h_{c_{II}}(R)$ (in blue) against the bond length (R) and the corresponding electron−electron van der Waals contributions (in green) for some noble-gas dimers: He_2, HeNe, and HeAr.

parameters control the electron correlation within the Hubbard model, small (large) U/t inducing dynamic (nondynamic) correlation. Hence, large U/t values prompt the electrons to distribute among the sites to minimize the electron repulsion. Figure 3 presents plots of the Coulomb hole at various values of U/t for the two-electron two-site Hubbard model in real space.[36] At low U/t values, the system is barely affected by correlation, thus dynamic correlation dominates (small $h_{c_I}(s)$ and large $h_{c_{II}}(s)$) and the electron pairs distribute equally between on-site and intersite components. As U/t grows, nondynamic correlation dominates and $h_{c_I}(s)$ becomes more important, being the prevailing contribution between sites.

He Series.[34] The He isoelectronic series is perhaps the simplest series of systems dominated by dynamic correction.[38] As the atomic number Z increases, the electron correlation of He(Z) tends to a constant and the exact electron density barely distinguishes from the HF one. In Figure 4 we observe that $h_{c_I}(s)$ decreases with the atomic number Z, and hence, $h_{c_{II}}$ completely takes over.

N Hydrogen Atoms.[34] The size consistency of our approach and its ability to measure spin entanglement is examined in

DOI: 10.1021/acs.jpclett.9b01376
J. Phys. Chem. Lett. 2019, 10, 4032−4037

Figure 5. We have plotted the Coulomb hole of the N-vertex polyhedron resulting from N hydrogen atoms separated by 10 Å from the center of the polyhedron. At these large separations, the hydrogen atoms only interact to each other through entanglement, and this is the only term that remains in the cumulant[4] (i.e., in $\Delta\rho_2^{c_{II}}$), which shows a linear behavior with N (see Figure 5). As in previous systems, $h_{c_{II}}$ is short ranged and its contribution to the energy grows linearly with N. These systems can be classified as nondynamic correlated because h_{c_I} is mostly long ranged and peaks at the same positions of the intracule density maxima. The planar D_{4h}/D_{2h} potential energy surface of H_4 has also been used for discriminating between dynamic and nondynamic correlation[39] and is given in the Supporting Information.

van der Waals (vdW) Interactions.[34] Figure 6 includes plots of the Coulomb hole of the helium dimer. h_c compares satisfactorily to earlier calculations.[40] The dynamic long-range interaction between the two noble-gas atoms is reflected by the second peak of the intracule density, whereas the interaction of the electron pair within each helium shows in the first peak. Regardless of the bond length, $h_{c_{II}}$ dominates, indicating that the correlation is dynamic and mainly affects the electron pair within each He. Unlike H_2, there is very little long-range nondynamic correlation in this system; however, at all distances, the long-range part of $h_{c_{II}}$ peaks around the bond-length distance (see the inset plots of Figure 6). The plot in Figure 7 presents $h_{c_{II}}(R)$ against the bond length, R, revealing a R^{-3} decay. It is a textbook fact that the pairwise vdW energy decays like R^{-6}.[41] Using perturbation theory, some of us have recently proved that the vdW contribution to $h_{c_{II}}$ should actually decay like R^{-3}, the integration of $h_{c_{II}}(s)/s$ over s yielding a fraction of the Coulombic interaction (eq 3) due to London dispersion forces and, therefore, decaying as R^{-6}.[28] Figure 7 includes plots for other noble-gas dimers, which also satisfy this property. Most density functional theory (DFT) practitioners add *ad hoc* empirical corrections to the energy for vdW interactions, and therefore, they only shift the relative energies of different conformers, yet the electronic structure of the system is not completely considered.[42] The present separation into correlation regimes unveils the target part of the pair density and the Coulomb hole, i.e., the long-range component of $h_{c_{II}}(s)$, which should be improved in order to incorporate the description of London dispersion forces and avoid the latter problem, thus opening a door to the accurate account of these forces within DFT and reduced density matrix functional theory (RDMFT).

In conclusion, eqs 8 and 9 represent a separation of the pair density and the Coulomb hole into components dominated by short- and long-range interactions. This result is expected to be important in the development of new hybrid electronic structure methods that can be employed in RDMFT[21,43] and other computational approaches. For instance, the HF reference in Figure 1 can be replaced by the Kohn–Sham system to adapt the present idea to DFT. It can be shown that the exchange-correlation functional can be entirely written in terms of the Kohn–Sham orbitals, $\Delta\rho_2^{c_I}$ and $\Delta\rho_2^{c_{II}}$. Hence, a template to construct density functional approximations, where the correlation components are treated separately, arises. Such possibility is already being explored in our laboratory.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.9b01376.

Analysis of the D_{2h}/D_{4h} H_4 planar molecule and the hydrogen molecule from ionic and covalent contributions; full plots of H_2, the two-site Hubbard dimer, the He isoelectronic series, and the N hydrogen atoms model (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ematito@gmail.com.

ORCID

Eloy Ramos-Cordoba: 0000-0002-6558-7821

Eduard Matito: 0000-0001-6895-4562

Author Contributions

‖These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank D. Casanova, X. Lopez, P. Salvador, and specially, P. M. W. Gill and J. M. Ugalde for helpful insights. This research has been funded by Spanish MINECO/FEDER Projects CTQ2014-52525-P, PGC2018-098212-B-C21, and EUIN2017-88605. We acknowledge doctoral grants BES-2015-072734 and FPU-2013/00176.

■ REFERENCES

(1) Löwdin, P.-O. 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. *Phys. Rev.* **1955**, *97*, 1474–1489.

(2) Cioslowski, J. Density-driven self-consistent-field method: Density-constrained correlation energies in the helium series. *Phys. Rev. A: At., Mol., Opt. Phys.* **1991**, *43*, 1223–1228.

(3) Gottlieb, A. D.; Mauser, N. J. New measure of electron correlation. *Phys. Rev. Lett.* **2005**, *95*, 123003.

(4) Raeber, A.; Mazziotti, D. A. Large eigenvalue of the cumulant part of the two-electron reduced density matrix as a measure of off-diagonal long-range order. *Phys. Rev. A: At., Mol., Opt. Phys.* **2015**, *92*, 052502.

(5) Benavides-Riveros, C. L.; Lathiotakis, N. N.; Schilling, C.; Marques, M. A. Relating correlation measures: The importance of the energy gap. *Phys. Rev. A: At., Mol., Opt. Phys.* **2017**, *95*, 032507.

(6) Valderrama, E.; Ludeña, E. V.; Hinze, J. Analysis of dynamical and nondynamical components of electron correlation energy by means of local-scaling density-functional theory. *J. Chem. Phys.* **1997**, *106*, 9227–9235.

(7) Valderrama, E.; Ludeña, E. V.; Hinze, J. Assessment of dynamical and nondynamical correlation energy components for the beryllium-atom isoelectronic sequence. *J. Chem. Phys.* **1999**, *110*, 2343–2353.

(8) Mok, D. K. W.; Neumann, R.; Handy, N. C. Dynamic and Nondynamic Correlation. *J. Phys. Chem.* **1996**, *100*, 6225–6230.

(9) Benavides-Riveros, C. L.; Lathiotakis, N. N.; Marques, M. A. Towards a formal definition of static and dynamic electronic correlations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 12655–12664.

(10) Ziesche, P. On relations between correlation, fluctuation and localization. *J. Mol. Struct.: THEOCHEM* **2000**, *527*, 35–50.

(11) Ramos-Cordoba, E.; Salvador, P.; Matito, E. Separation of dynamic and nondynamic correlation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24015–24023.

DOI: 10.1021/acs.jpclett.9b01376
J. Phys. Chem. Lett. 2019, 10, 4032–4037

(12) Ramos-Cordoba, E.; Matito, E. Local Descriptors of dynamic and nondynamic correlation. *J. Chem. Theory Comput.* **2017**, *13*, 2705−2711.
(13) Juhász, T.; Mazziotti, D. A. The cumulant two-particle reduced density matrix as a measure of electron correlation and entanglement. *J. Chem. Phys.* **2006**, *125*, 174105.
(14) Mazziotti, D. A. Approximate solution for electron correlation through the use of Schwinger probes. *Chem. Phys. Lett.* **1998**, *289*, 419−427.
(15) Ramos-Cordoba, E.; Salvador, P.; Piris, M.; Matito, E. Two new constraints for the cumulant matrix. *J. Chem. Phys.* **2014**, *141*, 234101.
(16) Cioslowski, J.; Piris, M.; Matito, E. Robust validation of approximate 1-matrix functionals with few-electron harmonium atoms. *J. Chem. Phys.* **2015**, *143*, 214101.
(17) Pastorczak, E.; Shen, J.; Hapka, M.; Piecuch, P.; Pernal, K. Intricacies of van der Waals interactions in systems with elongated bonds revealed by electron-groups embedding and high-level coupled-cluster approaches. *J. Chem. Theory Comput.* **2017**, *13*, 5404−5419.
(18) Savin, A. A combined density functional and configuration interaction method. *Int. J. Quantum Chem.* **1988**, *34*, 59−69.
(19) Grimme, S.; Waletzke, M. A combination of Kohn−Sham density functional theory and multi-reference configuration interaction methods. *J. Chem. Phys.* **1999**, *111*, 5645−5655.
(20) Bao, J. J.; Gagliardi, L.; Truhlar, D. G. Multiconfiguration pair-density functional theory for doublet excitation energies and excited state geometries: the excited states of CN. *Phys. Chem. Chem. Phys.* **2017**, *19*, 30089−30096.
(21) Piris, M. Global Method For The Electron Correlation. *Phys. Rev. Lett.* **2017**, *119*, 063002.
(22) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540−3544.
(23) Toulouse, J.; Gerber, I. C.; Jansen, G.; Savin, A.; Angyán, J. G. Adiabatic-connection fluctuation-dissipation density-functional theory based on range separation. *Phys. Rev. Lett.* **2009**, *102*, 096404.
(24) Toulouse, J.; Colonna, F.; Savin, A. Long-range short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A: At., Mol., Opt. Phys.* **2004**, *70*, 062505.
(25) Baer, R.; Livshits, E.; Salzner, U. Tuned range-separated hybrids in density functional theory. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85−109.
(26) Garrett, K.; Sosa Vazquez, X.; Egri, S. B.; Wilmer, J.; Johnson, L. E.; Robinson, B. H.; Isborn, C. M. Optimum exchange for calculation of excitation energies and hyperpolarizabilities of organic electro-optic chromophores. *J. Chem. Theory Comput.* **2014**, *10*, 3821−3831.
(27) Vuckovic, S.; Irons, T. J. P.; Wagner, L. O.; Teale, A. M.; Gori-Giorgi, P. Interpolated energy densities, correlation indicators and lower bounds from approximations to the strong coupling limit of DFT. *Phys. Chem. Chem. Phys.* **2017**, *19*, 6169−6183.
(28) Via-Nadal, M.; Rodríguez-Mayorga, M.; Matito, E. A Salient Signature of van der Waals Interactions. *Phys. Rev. A: At., Mol., Opt. Phys.* **2017**, *96*, 050501.
(29) Coulson, C. A.; Neilson, A. H. Electron correlation in the ground state of helium. *Proc. Phys. Soc., London* **1961**, *78*, 831−837.
(30) Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Via-Nadal, M.; Piris, M.; Matito, E. Comprehensive benchmarking of density matrix functional approximations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 24029−24041.
(31) Rigorously speaking, the condition that the coefficients should satisfy is that their N^{th} root, N being the number of electrons, is close to one for the system to be dominated by dynamic correlation. Otherwise, we would not correctly consider as such composites of noninteracting dynamic-correlation-driven systems, e.g., an infinite number of noninteracting helium atoms.
(32) Kutzelnigg, W.; Mukherjee, D. Cumulant expansion of the reduced density matrices. *J. Chem. Phys.* **1999**, *110*, 2800−2809.
(33) Ernzerhof, M.; Burke, K.; Perdew, J. P. Long-range asymptotic behavior of ground-state wave functions, one-matrices, and pair densities. *J. Chem. Phys.* **1996**, *105*, 2798−2803.
(34) We have performed full configuration interaction (FCI) calculations with the aug-cc-pVDZ basis set for He_2 and H_n (n = 2−8), and single and double configuration interactions calculations with the aug-cc-pVTZ basis set for HeNe and HeAr molecules. For the isoelectronic series of He, we have used an even-tempered basis set of 5s, 5p, and 5d functions, optimized following the procedure described elsewhere.[44]
(35) Mercero, J. M.; Valderrama, E.; Ugalde, J. M. In *Metal−Ligand Interactions*; Russo, N., Salahub, D. R., Witko, M., Eds.; Kluwer Academic Publishers: The Netherlands, 2003; pp 205−239.
(36) Carrascal, D.; Ferrer, J.; Smith, J. C.; Burke, K. The Hubbard dimer: a density functional case study of a many-body problem. *J. Phys.: Condens. Matter* **2015**, *27*, 393001.
(37) Deur, K.; Mazouin, L.; Fromager, E. Exact ensemble density functional theory for excited states in a model system: Investigating the weight dependence of the correlation energy. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95*, 035120.
(38) Chakravorty, S. J.; Gwaltney, S. R.; Davidson, E. R.; Parpia, F. A.; p Fischer, C. F. Ground-state correlation energies for atomic ions with 3 to 18 electrons. *Phys. Rev. A: At., Mol., Opt. Phys.* **1993**, *47*, 3649.
(39) Ramos-Cordoba, E.; Lopez, X.; Piris, M.; Matito, E. H_4: A challenging system for natural orbital functional approximations. *J. Chem. Phys.* **2015**, *143*, 164112.
(40) Piris, M.; Lopez, X.; Ugalde, J. Correlation holes for the helium dimer. *J. Chem. Phys.* **2008**, *128*, 134102.
(41) Pauling, L.; Wilson, E. B. *Introduction to Quantum Mechanics*; Dover Publications, Inc.: New York, 1935.
(42) Hermann, J.; DiStasio, R. A., Jr; Tkatchenko, A. First-Principles Models for van der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications. *Chem. Rev.* **2017**, *117*, 4714−4758.
(43) Piris, M.; Ugalde, J. M. Perspective on natural orbital functional theory. *Int. J. Quantum Chem.* **2014**, *114*, 1169−1175.
(44) Matito, E.; Cioslowski, J.; Vyboishchikov, S. F. Properties of harmonium atoms from FCI calculations: Calibration and benchmarks for the ground state of the two-electron species. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6712.

DOI: 10.1021/acs.jpclett.9b01376
J. Phys. Chem. Lett. 2019, 10, 4032−4037

Supplemental material: Singling Out Dynamic and Nondynamic Correlation

Mireia Via-Nadal[†],[1,2] Mauricio Rodríguez-Mayorga[†],[1,2,3]

Eloy Ramos-Cordoba,[1,2] and Eduard Matito[2,4]

[1]*Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Euskadi, Spain*

[2]*Donostia International Physics Center (DIPC), 20080 Donostia, Euskadi, Spain*

[3]*Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, University of Girona, C/ Maria Aurèlia Capmany, 69, 17003 Girona, Catalonia, Spain*

[4]*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Euskadi, Spain.*

(Dated: May 9, 2020)

SUPPLEMENTARY COULOMB HOLES

The Hydrogen molecule

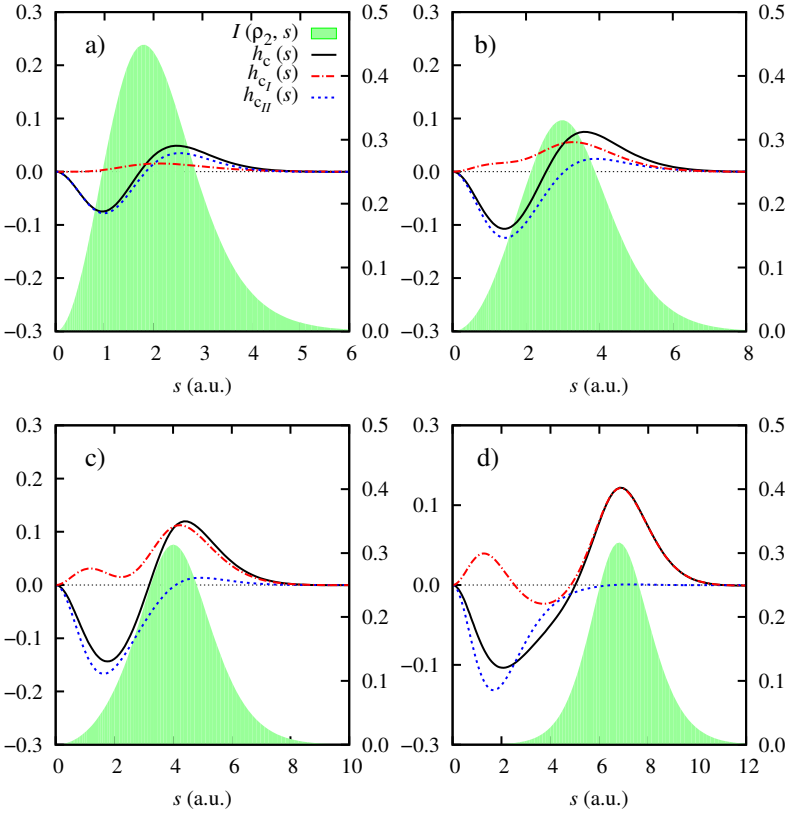


FIG. S1. The total Coulomb hole (black), $h_{c_{II}}$ (blue) and h_{c_I} (red) correlation components, and the exact (FCI) intracule density (shadowed green region, right y-axis) of the H_2 molecule at a) 1.32 , b) 2.83, c) 3.78, and d) 6.61 a.u.

The Helium dimer

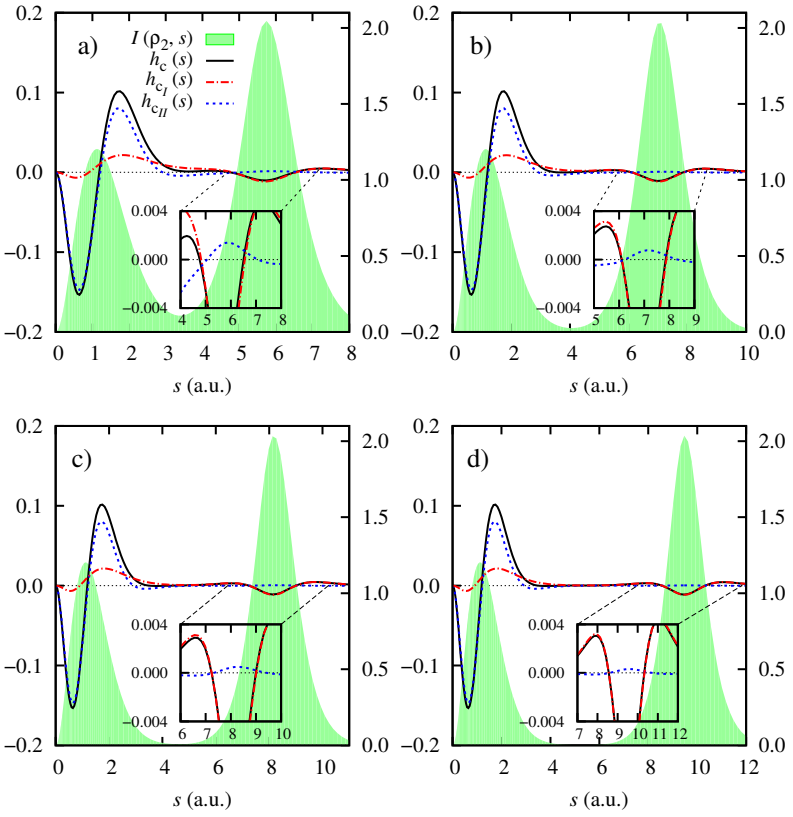


FIG. S2. The total Coulomb hole (black), $h_{c_{II}}$ (blue) and h_{c_I} (red) correlation components of the He_2 molecule at a) 5.67, b) 6.99, c) 8.13, and d) 9.45 a.u.. The inset plots reproduce the ones above on a narrower interval. The green-shadowed region represents the exact (FCI) intracule density.

The Hubbard model

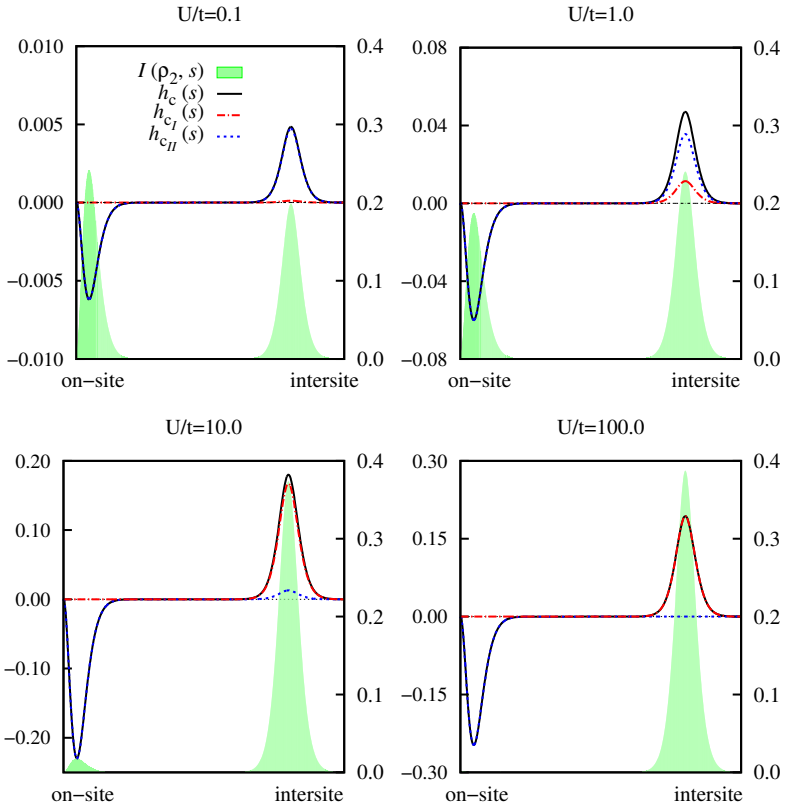


FIG. S3. The Coulomb hole (black), $h_{c_{II}}$ (blue) and h_{c_I} (red) correlation components of the half-filled two-site real-space Hubbard dimer for various U/t values. The green-shadowed region (right y-axis) corresponds to the exact (FCI) intracule density.

The He series

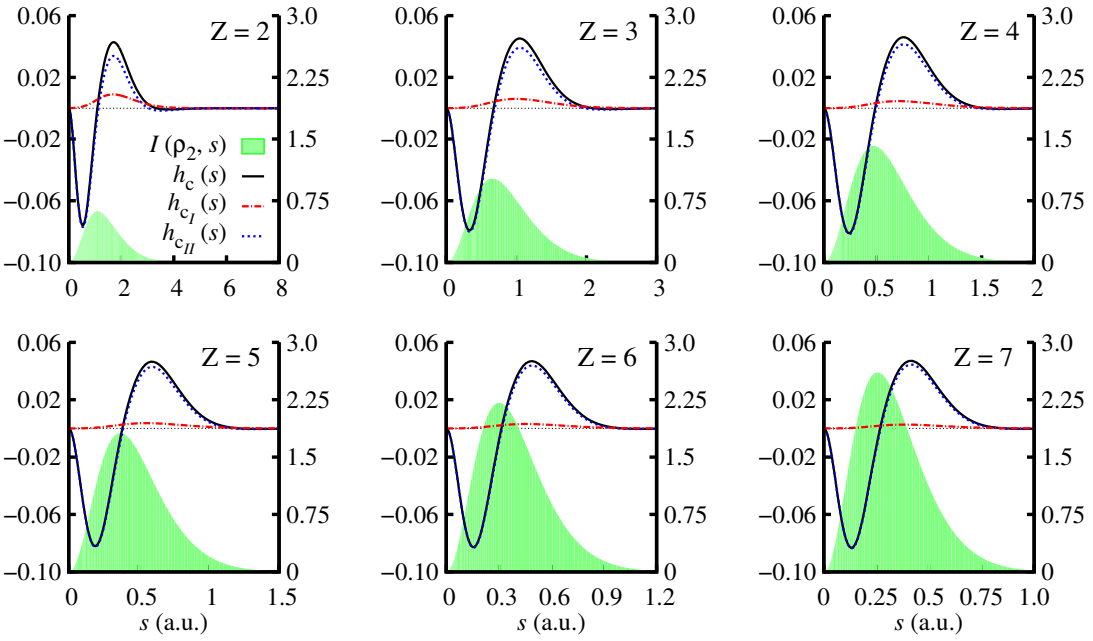


FIG. S4. The total Coulomb hole (black), $h_{c_{II}}$ (blue) and h_{c_I} (red) correlation components, and the exact (FCI) intracule density (shadowed green region, right y-axis) of the He isoelectronic series.

N Hydrogen atoms

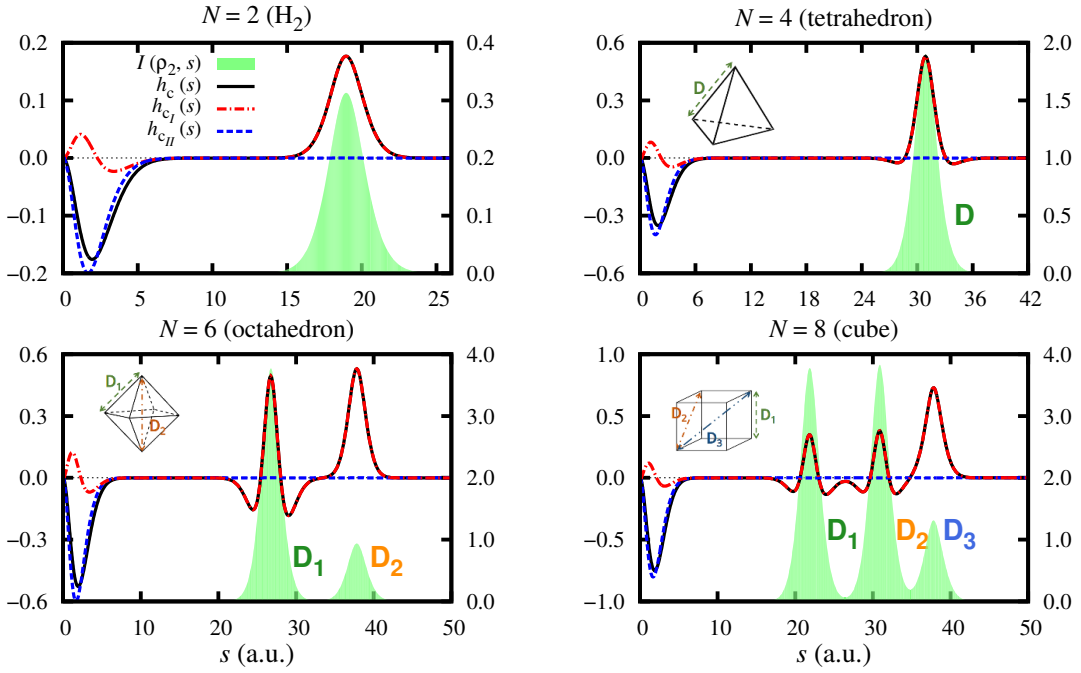


FIG. S5. The Coulomb hole (black), $h_{c_{II}}$ (blue) and h_{c_I} (red) correlation components of $N = 2, 4, 6$ and 8 hydrogen atoms, separated 10 Å from the geometric center of the polyhedra. The green-shadowed region (right y-axis) corresponds to the exact (FCI) intracule density. D, D_1, D_2 and D_3 indicate the different distances between the H atoms placed at the vertices of the polyhedra.

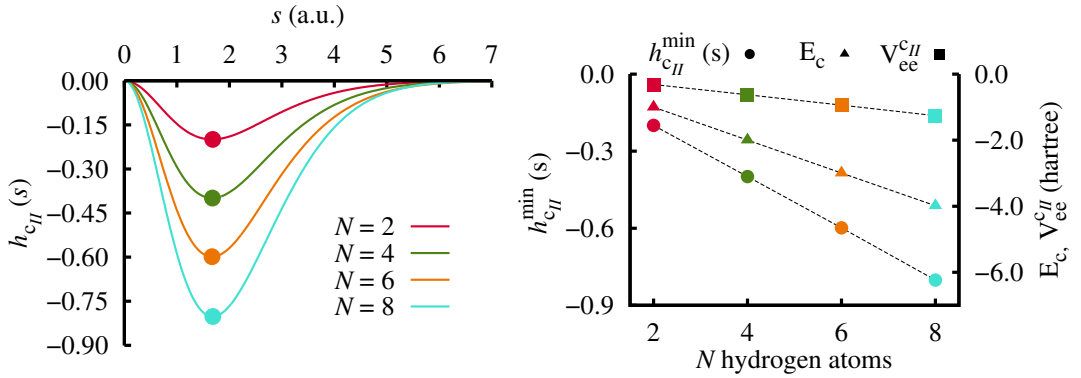


FIG. S6. The $h_{c_{II}}$ correlation component of the Coulomb hole of $N = 2, 4, 6$ and 8 hydrogen atoms separated 10 Å from the center of mass (see Fig. S5) The r.h.s. plot displays the minimal value of the $h_{c_{II}}^{min}$ component of the Coulomb hole, the part of electron-electron repulsion that corresponds to $\Delta\rho_2^{c_{II}}$, *i.e.* $V_{ee}^{c_{II}} = \int ds\ h_{c_{II}}(s)/s$, and the correlation energy (E_c) as a funcion of N.

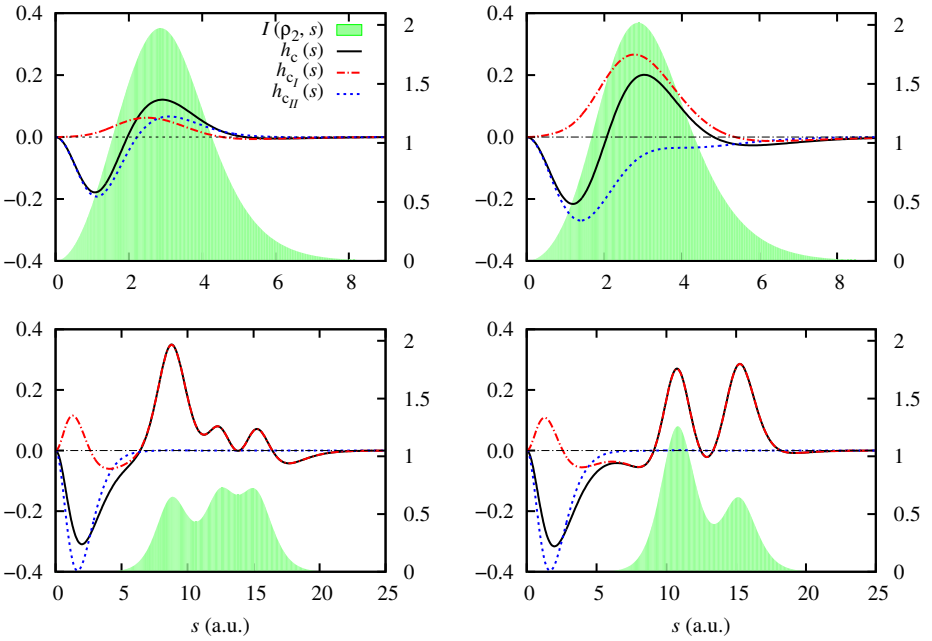


FIG. S7. The Coulomb hole (black), $h_{c_{II}}$ (blue) and h_{c_I} (red) correlation components of four conformations of D_{4h}/D_{2h} H_4, corresponding (from left to right and top to bottom) to $R = 1.51$ a.u. and $\theta = 0.39\pi$, $R = 1.51$ a.u. and $\theta = \pi/2$, $R = 7.56$ a.u. and $\theta = 0.39\pi$, and $R = 7.56$ a.u. and $\theta = \pi/2$, respectively. The green-shadowed region (right y-axis) corresponds to the exact (FCI) intracule density.

The planar D_{4h}/D_{2h} potential energy surface (PES) of H_4 [1] is determined by the distance from one hydrogen to the center of mass R and the angle between the center of mass and the two closest hydrogen atoms θ. This system is frequently used for discriminating different correlation types [2–4], being the regime of the PES dominated by nondynamic correlation characterized by large R or $\theta \approx \pi/2$, whereas the regime dominated by dynamic correlation is present at small R and $\theta \ll \pi/2$. The Coulomb hole contributions in the plots of Fig. S7 comply with the expected correlation trends. Indeed, h_{c_I} dominates for the square D_{4h} structure at large R. Despite the complicated mathematical structure of the long-range component of h_c (s), it is completely reproduced by h_{c_I} (s), which remarkably *only* depends on the 1-RDM.

H_2 MODEL

Let us analyze the contents of the pair density of the correlation components of the electron correlation, h_{c_I} and $h_{c_{II}}$, in terms of ionic and covalent contributions. We are going to use as a simple model H_2 described with a minimal basis set, namely, a normalized 1s orbital centered on each of the hydrogen atoms ($\chi_A(\mathbf{r})$ and $\chi_B(\mathbf{r})$). One can construct two orthonormal orbitals that are eigenfunctions of the Fock operator, a bonding

$$\varphi_+(\mathbf{r}) = \frac{1}{\sqrt{2(1+S_{AB})}}(\chi_A(\mathbf{r}) + \chi_B(\mathbf{r})), \tag{1}$$

and an antibonding orbital

$$\varphi_-(\mathbf{r}) = \frac{1}{\sqrt{2(1-S_{AB})}}(\chi_A(\mathbf{r}) - \chi_B(\mathbf{r})), \tag{2}$$

where S_{AB} is the overlap between atomic orbitals. The normalized HF ground state wavefunction is a Slater determinant built with the lower-in-energy-bonding molecular orbital as

$$\Phi_0(\mathbf{1},\mathbf{2}) = \frac{1}{\sqrt{2}}\varphi_+(\mathbf{r}_1)\varphi_+(\mathbf{r}_2)\left[\alpha(\omega_1)\beta(\omega_2) - \beta(\omega_1)\alpha(\omega_2)\right], \tag{3}$$

where $\mathbf{1} = \{\mathbf{r}_1, \omega_1\}$ refer to spatial and spin coordinates. Within this basis set, the exact (or FCI) wavefunction can be written as a linear combination of the HF ground state and the double excited Slater determinants

$$\Psi(\mathbf{1},\mathbf{2}) = c_0\Phi_0(\mathbf{1},\mathbf{2}) + c_1\Phi_1(\mathbf{1},\mathbf{2}), \tag{4}$$

where

$$c_0^2 + c_1^2 = 1, \tag{5}$$

and

$$\Phi_1(\mathbf{1},\mathbf{2}) = \frac{1}{\sqrt{2}}\varphi_-(\mathbf{r}_1)\varphi_-(\mathbf{r}_2)\left[\alpha(\omega_1)\beta(\omega_2) - \beta(\omega_1)\alpha(\omega_2)\right]. \tag{6}$$

As a prototype of a system dominated by nondynamic correlation, we are going to study this model at a large interatomic (R) distance. At this regime, the FCI solution is reached when $c_0 = -c_1 = \frac{1}{\sqrt{2}}$. Moreover, the product $\chi_A(\mathbf{r})\chi_B(\mathbf{r})$ tends to 0, and so does S_{AB}.

The exact pair density is written as

$$\begin{aligned}\rho_2(\mathbf{r}_1,\mathbf{r}_2)|_{R\to\infty} &= \int\int d\omega_1 d\omega_2 |\Psi(\mathbf{1},\mathbf{2})|^2|_{R\to\infty} \\ &= \frac{1}{2}(|\chi_A(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2)\end{aligned} \tag{7}$$

containing only covalent terms as expected. The exact 1-RDM (needed to construct the single-determinant approximation) is

$$\begin{aligned}\rho_1(\mathbf{r}_1';\mathbf{r}_1)|_{R\to\infty} &= \int\int d\omega_1 dx_2 \Psi^*(\mathbf{1}',\mathbf{2})\Psi(\mathbf{1},\mathbf{2})|_{\omega_1'=\omega_1} \\ &= \chi_A^*(\mathbf{r}_1')\chi_A(\mathbf{r}_1) + \chi_B^*(\mathbf{r}_1')\chi_B(\mathbf{r}_1).\end{aligned} \tag{8}$$

The HF pair density in terms of atomic orbitals at the strong correlation limit is

$$\begin{aligned}\rho_2^{\mathrm{HF}}(\mathbf{r}_1,\mathbf{r}_2)|_{R\to\infty} &= \int\int d\omega_1 d\omega_2 |\Phi_0(\mathbf{1},\mathbf{2})|^2|_{R\to\infty} \\ &= \frac{1}{4}[|\chi_A(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2 \\ &\quad +|\chi_A(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2],\end{aligned} \tag{9}$$

where the first two terms are the spurious ionic contributions and the last two correspond to half the covalent terms obtained with the exact pair density. The single determinant pair density is

$$\begin{aligned}\rho_2^{\mathrm{SD}}(\mathbf{r}_1,\mathbf{r}_2)|_{R\to\infty} &= \frac{1}{2}\rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2) - \frac{1}{4}\rho_1(\mathbf{r}_1;\mathbf{r}_2)\rho_1(\mathbf{r}_2;\mathbf{r}_1)|_{R\to\infty} \\ &= \frac{1}{2}(|\chi_A(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2) \\ &\quad +\frac{1}{4}(|\chi_A(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2),\end{aligned} \tag{10}$$

where the first two terms are covalent and the last two ionic. One can see that the covalent contribution is equal to the exact one, indicating that a single determinant pair density is

able to exactly describe the covalent (long-range in this case) contributions to the exact pair density. On the other hand, the spurious ionic terms are equivalent to the HF ones.

We define a pure ionic and a covalent pair density for a two-electron system as

$$\rho_2^{\text{ion}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{2}(|\chi_A(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2) \tag{11}$$

$$\rho_2^{\text{cov}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{2}(|\chi_A(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2). \tag{12}$$

The pair densities described above can be combined to define h_{c_I} and $h_{\text{c}_{II}}$ from the exact pair density. The pair density of the first type of correlation $\Delta\rho_2^{\text{c}_I}$ is

$$\begin{aligned}\Delta\rho_2^{\text{c}_I}(\mathbf{r}_1,\mathbf{r}_2)|_{R\to\infty} &= \frac{1}{4}(|\chi_A(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2) \\ &= \frac{1}{2}\rho_2^{\text{cov}}(\mathbf{r}_1,\mathbf{r}_2).\end{aligned} \tag{13}$$

$\Delta\rho_2^{\text{c}_I}(\mathbf{r}_1,\mathbf{r}_2)$ only contains covalent terms. By adding this fragment of the pair density to the HF one, the exact description of the covalent terms is recovered. The pair density for the *c*II type of correlation is then

$$\begin{aligned}\Delta\rho_2^{\text{c}_{II}}(\mathbf{r}_1,\mathbf{r}_2)|_{R\to\infty} &= -\frac{1}{4}(|\chi_A(\mathbf{r}_1)|^2|\chi_A(\mathbf{r}_2)|^2 + |\chi_B(\mathbf{r}_1)|^2|\chi_B(\mathbf{r}_2)|^2) \\ &= -\frac{1}{2}\rho_2^{\text{ion}}(\mathbf{r}_1,\mathbf{r}_2)\end{aligned} \tag{14}$$

It contains ionic terms that will cancel out when added to the reference HF pair density. Indeed, according to Fig. 2, $\Delta\rho_2^{\text{c}_{II}}$ removes the ionic contribution (*i.e.*, removes contributions that keep the electrons at short distances), whereas the $\Delta\rho_2^{\text{c}_I}$ adds the missing covalent contribution (*i.e.*, adds contributions that place one electron on each atom).

[1] We have performed full configuration interaction (FCI) calculations with the aug-cc-pVDZ basis set for planar H_4.

[2] E. Ramos-Cordoba and E. Matito, J. Chem. Theory Comput. **13**, 2705 (2017).

[3] K. Kowalski and K. Jankowski, Phys. Rev. Lett. **81**, 1195 (1998).

[4] M. P. E. Ramos-Cordoba, X. Lopez and E. Matito, J. Chem. Phys. **143**, 164112 (2015).

3.2 The separation of the c_I and c_{II} correlation components in the Coulomb hole, I

Range separation of the Coulomb hole[†]

Mireia Via-Nadal,[‡,¶] Mauricio Rodríguez-Mayorga,[‡] Eloy Ramos-Cordoba,[*,‡,¶] and Eduard Matito[*,‡,§]

‡*Donostia International Physics Center (DIPC), 20018 Donostia, Euskadi, Spain*

¶*Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Euskadi, Spain*

§*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Euskadi, Spain*

E-mail: eloy.raco@gmail.com; ematito@gmail.com

Abstract

A range-separation of the Coulomb hole into two components, one of them being predominant at short interelectronic separations ($h_{c_{II}}$) and the other at long distances (h_{c_I}), is exhaustively analyzed throughout various examples that put forward the most relevant features of this approach and how they can be used to develop efficient ways to capture electron correlation. We show that h_{c_I}, which only depends on the first-order reduced density matrix, can be used to identify molecules with a predominant nondynamic correlation regime, and differentiate between two types of nondynamic correlation, types A and B. Through the asymptotic properties of the hole components, we explain how h_{c_I} can retrieve the long-range part of electron correlation. We perform an exhaustive analysis of the hydrogen molecule in a minimal basis set, dissecting the hole contributions into spin components. We also analyze the simplest molecule presenting a dispersion interaction, and how $h_{c_{II}}$ helps to identify it. The study of several atoms in different spin states reveals that the Coulomb hole components permit to distinguish clear correlation regimes that are not apparent from the whole Coulomb hole. The results of

[†]This paper is dedicated to Paul Geerlings on occasion of his 70th birthday.

this work hold the promise to aid in the development of new electronic structure methods that efficiently capture electron correlation.

1 Introduction

In computational chemistry, the difference between the exact nonrelativistic electronic energy and the Restricted Hartree–Fock (RHF) one is known as correlation energy.[1] Even though the uncorrelated, self-consistent field (SCF) HF energy usually represents more than 98% of the total energy, the remaining is crucial to describe the chemistry of atoms and molecules (dissociation energies, reaction enthalpies, etcetera).[2,3] The description of the quantum system with the HF method lacks *electron correlation*, *i.e.*, a correct account of the correlated motion of the electrons. This conundrum is generally known as the many-body problem and it is one of main challenges in this field.[2,3] The study of *electron correlation* per se goes hand in hand with the development of electronic structure methods that do not bear a large computational cost.[4–22]

There exist many criteria to separate electron correlation, but the terminology involving non-dynamic (also static)[2,23,24] and dynamic[2,24] correlation is regularly used, as electronic structure methods are usually classified according to their ability to account for one of these types. Dynamic correlation is universally present in systems with more than two electrons, as it describes the motion of charged particles avoiding each other due to the electronic repulsion. Hence, this type of correlation increases with the number of electrons. A single-determinant picture along with a large number of low-contributing configurations is usually sufficient to portray this contribution. It is thus natural that the electron density displays very small changes with respect to the HF density.[15,22] Due to its nature, dynamic correlation affects electrons that are close to each other (short-ranged), but it is also responsible for non-covalent interactions such as the London dispersion forces (long-ranged).[25] Configuration interactions or coupled cluster with single and double excitations (CISD or CCSD),[26,27] Møller-Plesset second-order perturbation theory (MP2)[28] and density functional ap-

proximations (DFAs)[6,29] are methods that account for a large amount of dynamic correlation. On the other hand, nondynamic correlation is not universal, since it emerges when (near-)degeneracies are present in the system. It is characteristic of bond stretching, polyradical structures, entanglement, and high symmetries. The correct description of such correlation component requires a mix of largely-contributing configurations besides of the HF one.[30] Nondynamic correlation induces considerable changes in the electron density, caused by the mix of highly-contributing configurations in the CI vector.[14–16,20–22] Complete active space SCF (CASSCF),[31] density matrix renormalization group (DMRG),[32] and multi-configurational SCF (MCSCF)[33,34] are methods that are able to retrieve a fair amount of nondynamic correlation.

The ability to simultaneously tackle both types of correlation is present in very few electronic structure methods, such as the complete active space with second-order perturbation theory (CASPT2),[35,36] multireference configuration interactions (MRCI)[37,38] or, most recently, the adiabation-connection MCSCF (AC-MCSCF)[39] and the ΔNO.[40] Nevertheless, these methods are far from exact and the scaling of their computational cost with the number of electrons in the system represents a big drawback. Because of this, systems presenting both dynamic and nondynamic correlation types have become one of the greatest challenges in modern electronic structure methods. Latterly, an increasing interest in hybrid schemes such as range-separated methods[41,42] has appeared to confront the exposed problem. These methods aim to recover both correlation types by treating short- and long-range interactions with two different methodologies, according to their ability to recover one of the correlation components.[41–51] Although range-separation hybrid schemes provide a splitting of the pair density and the interelectronic coordinate, the separation is not custom-built for the correlation type present in the system.

Some studies have suggested measures of quantifying dynamic and nondynamic correlation,[8–19] some of them based on the correlation energy.[14–19] As reported by Coleman,[52] the use of electron-pair distribution functions to study electron correlation must lead to a useful understanding of both

short- and long-range interelectronic interactions. In line with the former statement, we have recently proposed a general decomposition scheme to separate the correlation part of the pair density into two components that permits the identification of systems with prevalent dynamic or nondynamic correlation.[20–22] From this scheme, scalar[20] and local[21] measures of dynamic and nondynamic correlation have been developed. Finally, a range-separation of the Coulomb hole has been introduced, providing the dominance of one component at short ranges (c_{II}), and the other one (c_I) at long ranges.[22]

In the present work, the range separation of the Coulomb hole is exhaustively analyzed throughout various examples that put forward the most relevant features of this approach and how they can be used to develop efficient ways to capture electron correlation. First of all, we introduce the range-separation of the Coulomb hole and its rationalization. Through the asymptotic properties of the hole parts, we explain how the component based only on the first-order reduced density matrix can retrieve the long-range part of electron correlation. Second, we perform an exhaustive analysis of the hydrogen molecule in a minimal basis set, dissecting the hole contributions into spin components. Third, we analyze the simplest molecule presenting a dispersion interaction, and how one of the Coulomb hole components helps identifying it. This dispersion signature is also identified in the remaining molecules of the manuscript, highlighting its universal character. Afterwards, we analyze the Coulomb holes of several atoms in different spin states, finding that the Coulomb hole components permit to distinguish clear correlation regimes that are not apparent from the whole Coulomb hole. Finally, we analyze the two types of nondynamic correlation, types A and B,[11] and show that they can be both captured by the first component of the Coulomb hole.

The results of this work hold the promise to aid in the development of new electronic structure methods that efficiently capture electron correlation. In particular, the models of the c_{II} component can be straightforwardly used in the reduced density matrix functional theory,[53–56] although they are not limited to this theory.

2 Methodology

The pair density of a N-electron system described by a $\Psi(\mathbf{1},\ldots,\boldsymbol{N})$ wavefunction is

$$\rho_2(\mathbf{1},\mathbf{2}) = N(N-1)\int \mathrm{d}_{\mathbf{3}}\ldots\mathrm{d}_{\boldsymbol{N}}|\Psi(\mathbf{1},\ldots,\boldsymbol{N})|^2, \tag{1}$$

where we have assumed McWeeny's normalization,[57] which accounts for the number of electron pairs in the system, and the variables $(\mathbf{1},\mathbf{2},\ldots)$ refer to both space and spin coordinates, $\mathbf{1} \equiv \vec{r}_1, \sigma_1$. The pair density contains information about the probability of finding electron 1 at position and spin $\mathbf{1}$ and electron 2 at position and spin $\mathbf{2}$ when the remaining of $N-2$ electrons are placed anywhere. The analysis of the pair density is arduous since it depends on six space and two spin variables. Conversely, the radial intracule density is a reduced form of the pair density that only depends on the interelectronic distance or range separation, s, and still retains the necessary information to calculate the electronic repulsion energy,

$$V_{ee} = \frac{1}{2}\int \mathrm{d}s \frac{I(\rho_2, s)}{s}, \tag{2}$$

being

$$I(\rho_2, s) = \int \mathrm{d}_{\mathbf{1}}\mathrm{d}_{\mathbf{2}}\,\rho_2(\mathbf{1},\mathbf{2})\delta(s - r_{12}), \tag{3}$$

where r_{12} is the Euclidean distance between electrons 1 and 2, and $\delta(s-r_{12})$ is the Dirac delta. Eq 3 is the radial intracule density and it corresponds to the probability distribution of interelectronic distances s between two electrons. It provides information about the electron-pair relative motion within atoms and molecules. This distribution can also be obtained experimentally from X-ray scattering cross-sections.[58–61]

Analogously to Löwdin's definition of correlation energy,[1] Coulson defined the correlated pair density as the difference between the exact and an uncorrelated pair density, which was chosen to

be the Hartree–Fock (HF) one,[62]

$$\Delta\rho_2^{\mathrm{c}}(\mathbf{1},\mathbf{2}) = \rho_2(\mathbf{1},\mathbf{2}) - \rho_2^{\mathrm{HF}}(\mathbf{1},\mathbf{2}). \tag{4}$$

The intracule of the correlation pair density $\Delta\rho_2^{\mathrm{c}}(\mathbf{1},\mathbf{2})$ is known as Coulson's Coulomb hole:[63–65]

$$h_{\mathrm{c}}(s) = I(\Delta\rho_2^{\mathrm{c}}, s) = I(\rho_2, s) - I\left(\rho_2^{\mathrm{HF}}, s\right). \tag{5}$$

The Coulomb hole is thus a probability density difference that reflects the effect of switching from the mean-field HF approximation, which underestimates the electronic repulsion, to a correlated framework. This translates into an increase in the expected interelectronic distance vector, $\langle s_{\mathrm{HF}}\rangle < \langle s\rangle$, caused by the correct account of the electronic repulsion ($\langle s\rangle = \int I(s) s ds$).

We have introduced elsewhere[20–22] a splitting of the correlated pair density, eq 4, using the single-determinant approximation to the pair density,[1]

$$\rho_2^{\mathrm{SD}}(\rho_1, \mathbf{1}, \mathbf{2}) = \rho(\mathbf{1})\rho(\mathbf{2}) - |\rho_1(\mathbf{1};\mathbf{2})|^2, \tag{6}$$

where $\rho_1(\mathbf{1};\mathbf{2})$ is the first-order reduced density matrix (1-RDM),

$$\rho_1(\mathbf{1};\mathbf{1}') = \int \mathrm{d}_{\mathbf{2}} \ldots \mathrm{d}_{\boldsymbol{N}} \Psi(\mathbf{1},\mathbf{2},\ldots,\boldsymbol{N})\, \Psi^*(\mathbf{1}',\mathbf{2},\ldots,\boldsymbol{N}), \tag{7}$$

and its diagonal part, $\rho(\mathbf{1}) = \rho(\mathbf{1},\mathbf{1})$, is the electron density. The SD ansatz, eq 6, takes advantage of the HF expression for the pair density but uses the actual 1-RDM to generate an approximation to the pair density. Obviously, it returns the HF pair density when the HF 1-RDM is used, *i.e.*,

$$\rho_2^{\mathrm{HF}}(\mathbf{1},\mathbf{2}) = \rho_2^{\mathrm{SD}}\left(\rho_1^{\mathrm{HF}}, \mathbf{1}, \mathbf{2}\right). \tag{8}$$

The SD approximation includes neither short-[66] nor long-range[67] dynamic correlation, that is, it

only accounts for nondynamic correlation. Even though it cannot be guaranteed that *all* nondynamic correlation in the system is accounted for, it is legitimate to claim that the SD pair density, eq 6, includes some extent of it. In fact, the SD approximation captures most long-range electron correlation effects and usually presents a very small short-range contribution, except for systems with type B nondynamic correlation[11] that arises due to the relative degeneracy of its frontier orbitals (vide infra). Unlike the HF pair density, the SD approximation is in general not N-representable by construction, *i.e.*, this approximate pair density does not necessarily correspond to an N-particle fermionic wavefunction. The violation of the N-representability conditions may lead to spurious energies[68] and affect density matrix properties as the trace, $\mathrm{Tr}\left[\rho_2^{\mathrm{SD}}\right]$, or the positivity of its eigenvalues, among others.[66,69–73] Despite it is not N-representable, $\rho_2^{\mathrm{SD}}(\mathbf{1},\mathbf{2})$ is simply used in this context as a mathematical approximation to seize nondynamic correlation. Indeed, we have used it to separate eq 4 into two correlation contributions, $\Delta\rho_2^{c_I}$ and $\Delta\rho_2^{c_{II}}$,[22]

$$\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2}) = \rho_2^{\mathrm{SD}}(\rho_1,\mathbf{1},\mathbf{2}) - \rho_2^{\mathrm{SD}}\left(\rho_1^{\mathrm{HF}},\mathbf{1},\mathbf{2}\right) \tag{9}$$

$$\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2}) = \rho_2(\rho_1,\mathbf{1},\mathbf{2}) - \rho_2^{\mathrm{SD}}(\rho_1,\mathbf{1},\mathbf{2}), \tag{10}$$

which, along with the HF pair density, recover the exact pair density:

$$\rho_2(\mathbf{1},\mathbf{2}) = \rho_2^{\mathrm{HF}}(\mathbf{1},\mathbf{2}) + \Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2}) + \Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2}). \tag{11}$$

The partition of the pair density into the HF reference plus the c_I component and the c_{II} part is known as the Lieb partitioning of the pair density.[74–76] $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ is also known as the cumulant of the density matrix, which captures the correlation lacking in the 2-RDM.[77,78] $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ only depends on the actual 1-RDM and the HF one. $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ is actually a measure of the dissimilarity between the latter two matrices. Systems with nondynamic correlation present larger changes on the electron density with respect to HF than systems with dynamic correlation.[15] Therefore, $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ can be also regarded as a function measuring nondynamic correlation. In this sense, the information about dynamic correlation is expected to be fully contained in $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$, which

might also include nondynamic correlation to some extent.

A more explicit expression for eqs 9 and 10 can be casted:

$$\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2}) = |\rho_1^{\mathrm{HF}}(\mathbf{1};\mathbf{2})|^2 - |\rho_1(\mathbf{1};\mathbf{2})|^2 + \rho(\mathbf{1})\rho(\mathbf{2}) - \rho^{\mathrm{HF}}(\mathbf{1})\rho^{\mathrm{HF}}(\mathbf{2}) \quad (12)$$

$$\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2}) = \rho_2(\mathbf{1},\mathbf{2}) - \rho(\mathbf{1})\rho(\mathbf{2}) + |\rho_1(\mathbf{1};\mathbf{2})|^2. \quad (13)$$

If we employ the correlation components of the pair density, we can split the Coulomb hole into two hole components,

$$h_{\mathrm{c}}(s) = h_{\mathrm{c}_I}(s) + h_{\mathrm{c}_{II}}(s) = I\left(\Delta\rho_2^{c_I}, s\right) + I\left(\Delta\rho_2^{c_{II}}, s\right), \quad (14)$$

which permit the analysis of electron correlation in terms of interelectronic ranges.[22] The asymptotic properties of $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ determine the long-range behavior of the Coulomb hole components. The first important property of $h_{\mathrm{c}_I}(s)$ and $h_{\mathrm{c}_{II}}(s)$ is that they vanish for very large values of s as long as such large value of s corresponds to points $\mathbf{1}$ or $\mathbf{2}$ in $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ that are far from the molecule. Since both the HF and the exact pair density are zero for such points, one only needs to prove the same for the SD approximation. The latter is formed by two terms, the first one involving the product of two densities and, therefore, it vanishes in regions far from the molecule. The long-range asymptotics of the 1-RDM[79] also guarantees that the second term, including the square of a 1-RDM, vanishes under this condition. Second, we study the behavior of the two Coulomb hole components at short ranges. By virtue of the Pauli principle, the same-spin component of $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ vanishes when $r_{12} = 0$. Hence, one can easily prove that at the electronic coalescence points, $\Delta\rho_2^{c_I}(\vec{r}_1,\vec{r}_1) = 2\left(\rho^{\alpha}(\vec{r}_1)\rho^{\beta}(\vec{r}_1) - \rho^{\mathrm{HF},\alpha}(\vec{r}_1)\rho^{\mathrm{HF},\beta}(\vec{r}_1)\right)$. Since HF underestimates the electron-nucleus cusp, this quantity is greater than zero in the vicinity of the nuclei. As a matter of fact, our experience indicates that $h_{\mathrm{c}_I}(s)$ is usually greater than zero.[1] Conversely,

[1] A remarkable exception to this rule occurs in molecules that are dissociated by HF into fragments with a wrong number of electrons. See the last two examples given in this paper.

$h_{c_{II}}(s)$ is usually negative at short ranges because $\Delta\rho_2^{c_{II}}(\vec{r}_1,\vec{r}_1) = 2\left(\rho_2^{\alpha\beta}(\vec{r}_1,\vec{r}_1) - \rho^{\alpha}(r_1)\rho^{\beta}(r_1)\right)$ is negative at points close to the nuclei, which are the points that contribute the most to $\Delta\rho_2^{c_{II}}(\vec{r}_1,\vec{r}_1)$. In the following, we analyze $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ at long ranges to prove its short-ranged character. The long-range contributions come mostly from points which are very close to nuclei, typically points that are centered into two different atoms. For the sake of simplicity, let us choose the hydrogen molecule dissociation to illustrate this point. We can take the leading term in the expansion of $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ around the two electron-nucleus cusps, $\Delta\rho_2^{c_{II}}(R_{\mathrm{H}},R_{\mathrm{H}'}) = \rho_2(R_{\mathrm{H}},R_{\mathrm{H}'}) - \rho_2^{\mathrm{SD}}(R_{\mathrm{H}},R_{\mathrm{H}'})$, which we have fully developed in the Supporting Information. In Figure 1, we represent the ratio $\rho_2^{\mathrm{SD}}(R_{\mathrm{H}},R_{\mathrm{H}'})/\rho_2(R_{\mathrm{H}},R_{\mathrm{H}'})$ against the interatomic separation $R_{\mathrm{HH}'}$. The ratio is always greater than 0.8 and easily achieves 1.0 as the bond stretches, indicating that $\Delta\rho_2^{c_{II}}(R_{\mathrm{H}},R_{\mathrm{H}'})$ quickly vanishes at long distances and, therefore, $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ is expected to vanish quickly with r_{12} and, hence, be predominantly short-ranged.

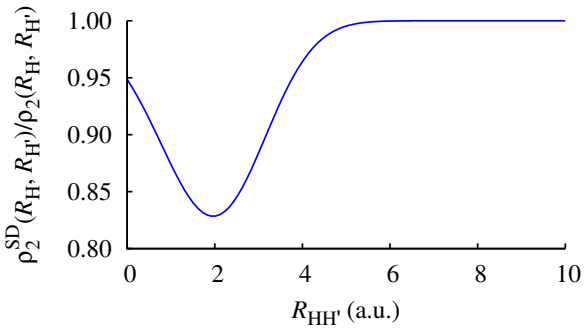


Figure 1: $\rho_2^{\mathrm{SD}}(R_{\mathrm{H}},R_{\mathrm{H}'})/\rho_2(R_{\mathrm{H}},R_{\mathrm{H}'})$ against the interatomic separation $R_{\mathrm{HH}'}$ for the minimal basis set calculation of the hydrogen molecule.

Finally, let us mention that the evaluation of the c_{II} component of the Coulomb hole at the interfragment separation distance R, $h_{c_{II}}(R)$, decays like R^{-3} when $R \to \infty$.[67] This dependency is connected with the well-known R^{-6} decay of the van der Waals (vdW) dispersion energy. $h_{c_{II}}(s)$ actually presents a narrow positive region at interelectronic distances around R that reflects the binding character of dispersion interactions, as we shall see in the examples covered in the Results section. Because dispersion interactions are weak in nature, the region in $h_{c_{II}}$ is proportionally

small.

3 Computational Details

Full configuration interactions (FCI) calculations have been run with a modified version of Knowles and Handy's program.[80,81] The Dunning's augmented correlation-consistent double zeta basis set (aug-cc-pVDZ)[82,83] is used, unless otherwise specified. For the Be isoelectronic series we have used the basis sets developed in a previous paper.[20,81] The density matrices and the intracule probability distributions have been obtained with the in-house DMN[84] and RHO2_OPS[85] codes, respectively, the latter using the algorithm of Cioslowski and Liu.[86]

4 Results and discussion

4.1 H_2 in minimal basis

Due to its simplicity, H_2 in a minimal basis (STO-3G)[87] becomes a perfect model to understand the partition of the Coulomb hole. Let us consider the equilibrium and a stretched geometry. In both cases, the total Coulomb hole is negative at the short range, indicating that HF overestimates the number of electron pairs at short interelectronic distances (see Figure 2). The hole of H_2 at equilibrium becomes positive for values of s larger than the bond length, as a consequence of HF's underestimation of number of electrons pairs at these interelectronic distances. The c_I part of the hole is positive and rather small, whereas c_{II} accounts for almost the entire shape of the total Coulomb hole. This is a clear indication that the role of nondynamic correlation is small at equilibrium. At the stretched geometry, the magnitude of the Coulomb hole is considerably larger and the maximum of the hole coincides with maximum of the intracule of the pair density. In this case, HF does a worse job in describing the distribution of electron pairs, providing even larger errors on both the short- and long-range components of the Coulomb hole. As a result of the separability

problem, which originates from the impossibility to separate two electrons occupying the same orbital on a restricted single-determinant wavefunction,[2] HF clearly underestimates the number of electrons pairs formed at large electron-electron separations and overestimates the number of pairs formed at short separations (see Figure 3). Interestingly, the two components of the Coulomb hole provide a clear-cut separation of both phenomena: c_I corresponds to the electron pairs missed by HF at long ranges and c_{II} corrects the overestimation of electron pairing at short range. Since the Coulomb hole is actually the correction to HF's two-electron distribution, its energetic contribution to the total energy can be regarded as the correction to the HF electron-electron repulsion. In this sense, both hole components contribute to the energy correction. However, the c_I contribution is small and comes entirely from the short-range part (which is not apparent in Figure 2, unless we plot $h_{c_I}(s)/s$) because $h_{c_I}(s)/s$ tends to zero for large s as we reach the dissociation limit. Conversely, other properties based on the pair density, such as the covariance of the electron populations of the two atoms,[88] are affected by the value of c_I at all ranges and, therefore, the correct account of h_{c_I} at short ranges would not be enough.

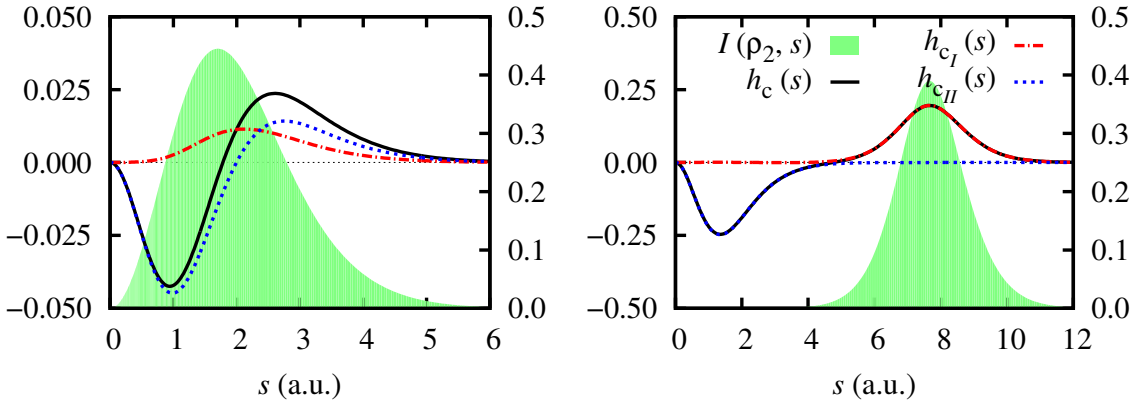


Figure 2: Coulomb hole (black), $h_{c_I}(s)$ (red) and $h_{c_{II}}(s)$ (blue) correlation components, and the intracule density (shadowed green region, right y-axis) of H_2 calculated with STO-3G basis set at the equilibrium distance (1.39 a.u. bond distance) and 7.56 a.u. bond distance.

Let us now dissect the hole and its components into spin contributions for the stretched geometry (see Figure 4). The ground state of H_2 has one electron of each spin, thus the same-spin

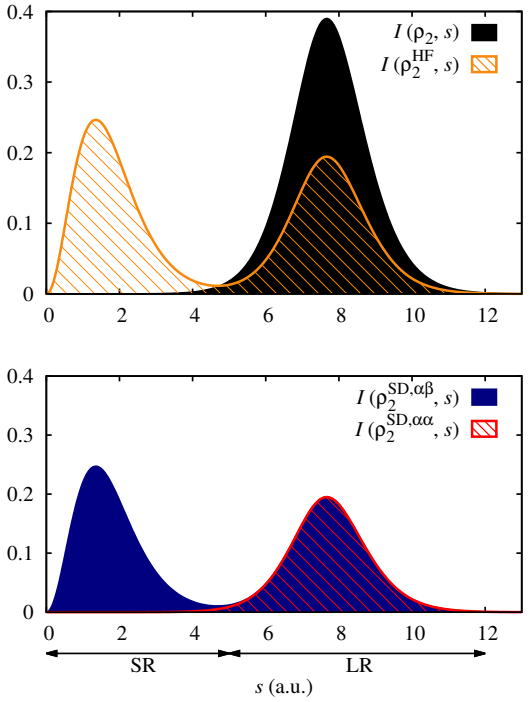


Figure 3: The intracule probability densities of H_2 at $R = 7.56$ a.u. bond distance calculated with the STO-3G basis set. Top: the exact (filled black) and the Hartree–Fock (dashed orange) intracules. Bot: the opposite- (filled blue) and same-spin (dashed red) components of the Single-Determinant approximation intracule. The filling has been included as a visualization aid.

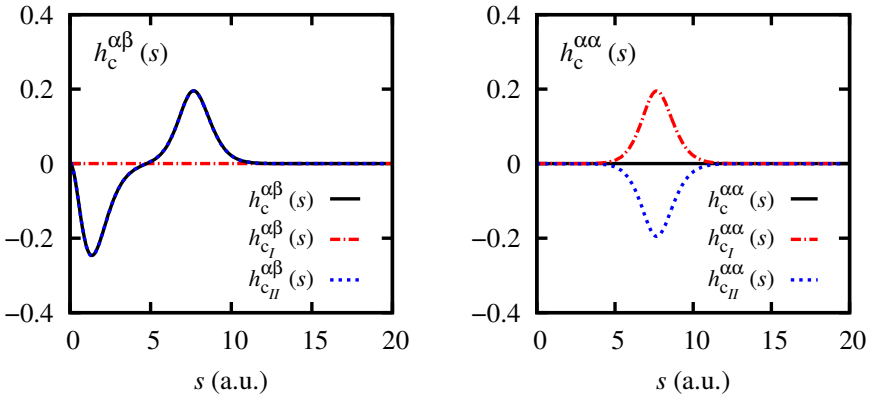


Figure 4: The opposite-(left) and same-spin (right) components of the Coulomb hole (h_c), $h_{c_I}(s)$ (red) and $h_{c_{II}}(s)$ (blue) correlation components of H_2 at $R = 7.56$ a.u. bond length calculated with the STO-3G basis set.

contribution of the Coulomb hole is zero, $h_{\mathrm{c}}^{\alpha\alpha}(s) = 0$. However, the same-spin contribution of the correlation components of the Coulomb hole, $h_{\mathrm{c}_I}^{\alpha\alpha}(s)$ and $h_{\mathrm{c}_{II}}^{\alpha\alpha}(s)$, is not zero. This reveals that the Coulomb hole components do not have any physical meaning; they are simply mathematical objects that are defined for convenience. Indeed, the so-defined $h_{\mathrm{c}_I}(s)$ of H_2 at dissociation arises uniquely from $h_{\mathrm{c}_I}^{\alpha\alpha}(s)$, which completely reproduces the long-range behavior of the total Coulomb hole. Conversely, $h_{\mathrm{c}_I}^{\alpha\beta}(s)$ is very small by definition, because neither HF nor SD pair densities explicitly introduce opposite-spin correlation and the difference between HF and the exact density is expected to be small (in the present case, where we employ a minimal basis set, the density difference is zero in the limit of H_2 dissociation). The long-range opposite-spin correlation is indirectly introduced in the c_I component through the addition of the $h_{\mathrm{c}_I}^{\alpha\alpha}(s)$ contribution. Obviously, for systems of only one alpha electron, $h_{\mathrm{c}_I}^{\alpha\alpha}(s) = -h_{\mathrm{c}_{II}}^{\alpha\alpha}(s)$ and, therefore, $h_{\mathrm{c}_{II}}^{\alpha\alpha}(s)$ is compensating for the long-range contribution of $h_{\mathrm{c}_{II}}^{\alpha\beta}(s)$, making $h_{\mathrm{c}_{II}}(s)$ short-ranged. In other words, $h_{c_I}(s)$ captures the long-range component of the Coulomb hole through the intracule of $\rho_2^{\mathrm{SD},\alpha\alpha}(\mathbf{1},\mathbf{2})$. In particular, the latter contributes a quadratic term on r_{12} at short ranges and its long-range contribution is dominated by the $\rho^\alpha(\mathbf{1})\rho^\alpha(\mathbf{2})$ term in $\rho_2^{\mathrm{SD},\alpha\alpha}(\mathbf{1},\mathbf{2})$. Hence, as first approximation to the long-range part of $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ one could take $\rho^\alpha(\mathbf{1})\rho^\alpha(\mathbf{2}) + \rho^\beta(\mathbf{1})\rho^\beta(\mathbf{2})$, which captures the long-range asymptotics of the Coulomb hole.

4.2 H_2 triplet: dispersion interactions

The triplet state of H_2 ($^3\Sigma_u^+$) is solely composed of two electrons with the same spin and constitutes probably the simplest model to study dispersion interactions.[89] In contrast to the ground state singlet ($^1\Sigma_u^+$), the triplet state $^3\Sigma_u^+$ (with spin projection either $M_s = 1$ or $M_s = -1$) is qualitatively well described by the HF configuration at any interatomic distance and bears the correct dissociation.[90] Indeed, compared to the singlet, the Coulomb hole is very small (see the left y-axis in Figure 5). This case does not present electron entanglement, therefore, nondynamic correlation does not arise when the two fragments are separated.

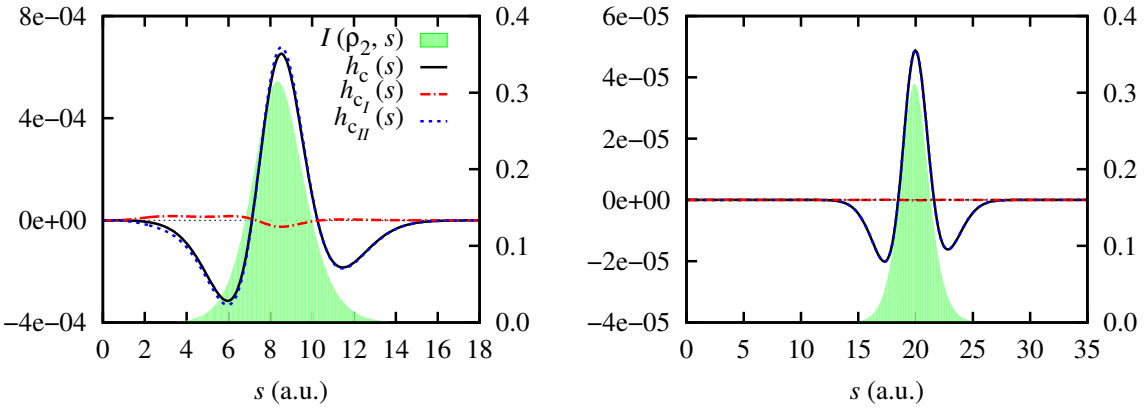


Figure 5: Coulomb hole (black), $h_{c_I}(s)$ (red) and $h_{c_{II}}(s)$ (blue) correlation components, and the intracule density (shadowed green region, right y-axis) of the $^3\Sigma_u^+$ state of H_2 molecule at 8.18 (left) and 19.84 a.u. (right) of bond distance.

Since we have employed a basis set including p orbitals, long-range dispersion interactions are taken into account.[91] Hence, the only peak present in the plots corresponds to dispersion interactions, *i.e.* dynamic correlation and, thus, it is captured by $h_{c_{II}}(s)$.[67] $h_{c_{II}}(s)$ peaks around the interatomic separation, $R_{HH'}$, and displays a shape that indicates that HF underestimates the number of electrons pairs that are separated $R_{HH'}$, which are placed at shorter and longer distances. On the other hand, there is almost no $h_{c_I}(s)$ contribution at equilibrium, and it is completely zero at the stretched geometry. The reason stems from the fact that $\rho_2^{SD}(\mathbf{1},\mathbf{2})$ is very close to $\rho_2^{HF}(\mathbf{1},\mathbf{2})$ due to the very small contribution of the electron correlation to the changes of the density or the first-order density. As we have recently proved,[22,67] dispersion interactions are characterized by a universal feature: $h_{c_{II}}(R_{HH'})$ decays as $R_{HH'}^{-3}$ for large $R_{HH'}$.

4.3 He-Ne atomic series.

Since correlation increases with the number of electrons, N, the Coulomb hole dimension is expected to increase with the atomic number, Z. The magnitude of the hole increases and shrinks with Z ($\langle s \rangle$ decreases due to larger attractive nucleus potential); however, this increase is not monotonic (see the plot for h_c in Figure S4). The splitting of the hole into the two correlation components

reveals that $h_{c_{II}}(s)$ (which includes the dynamic correlation effects of the hole) presents a systematic growth with Z (see the magnitude increase in the top plot of Figure 6). Instead, $h_{c_I}(s)$ reveals characteristic shapes according to the nature of each atom, in agreement with the expected nondynamic correlation behavior in these atoms (bottom plot in Figure 6).

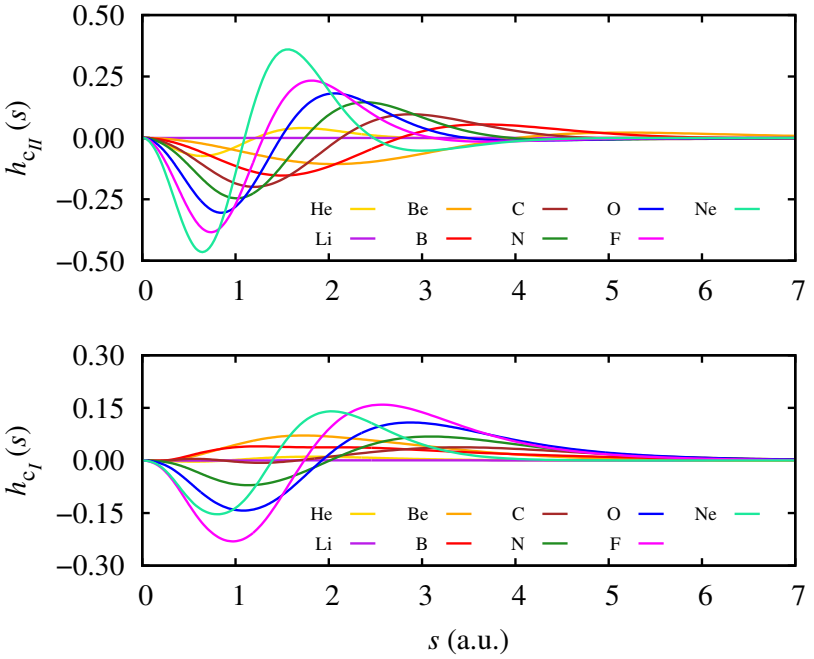


Figure 6: $h_{c_{II}}$ (top) and h_{c_I} (bot) correlation components of the Coulomb hole of the He-Ne atomic series in their lowest-lying multiplicities.

We have studied different multiplicities for carbon, nitrogen and oxygen: their lowest-lying states (highest spin multiplicity) and the states with the lowest multiplicity. Their holes are compiled in Figure 7. The most outstanding feature is that the Coulomb hole of all the atoms looks practically identical regardless the multiplicity and, therefore, one expects a similar correlation contribution to the electron-electron repulsion. Hence, a mere evaluation on the shape of the Coulomb hole does not provide insight about how electron correlation changes in different multiplicity states. However, the Coulomb hole components provide more information about electron correlation, since $h_{c_I}(s)$ and $h_{c_{II}}(s)$ have different shapes according to the spin state.

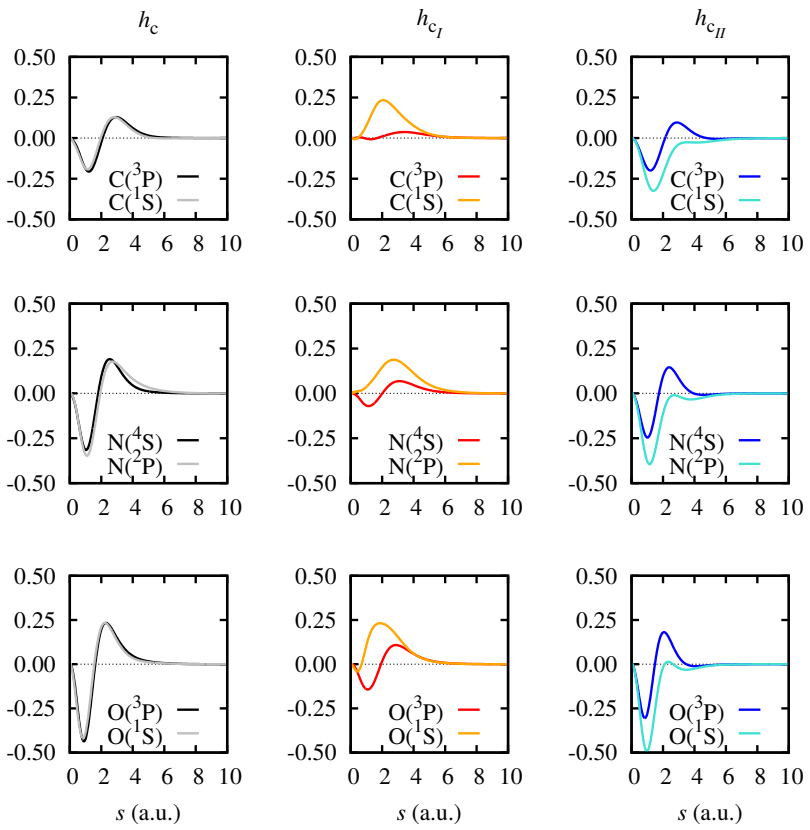


Figure 7: Coulomb hole $h_c(s)$, $h_{c_I}(s)$ and $h_{c_{II}}(s)$ correlation components of carbon, nitrogen and oxygen in both their ground state (black, red and blue, respectively) and minimum multiplicity state (gray, orange and turquoise).

More precisely, each state presents a characteristic profile, independently of the atom considered. In the ground states (highest spin multiplicities), both c_I and c_{II} components present the conventional hole shape, being negative at short interelectronic distances and positive at large ones. $h_{c_{II}}$ is, in general, larger than h_{c_I} and both increase systematically with the number of electrons, N (see Figure 7), in line with the fact that the HF description becomes less accurate as N increases in this set of atoms. Conversely, the atoms in their lowest spin multiplicity present mostly positive h_{c_I} and short-ranged negative $h_{c_{II}}$.

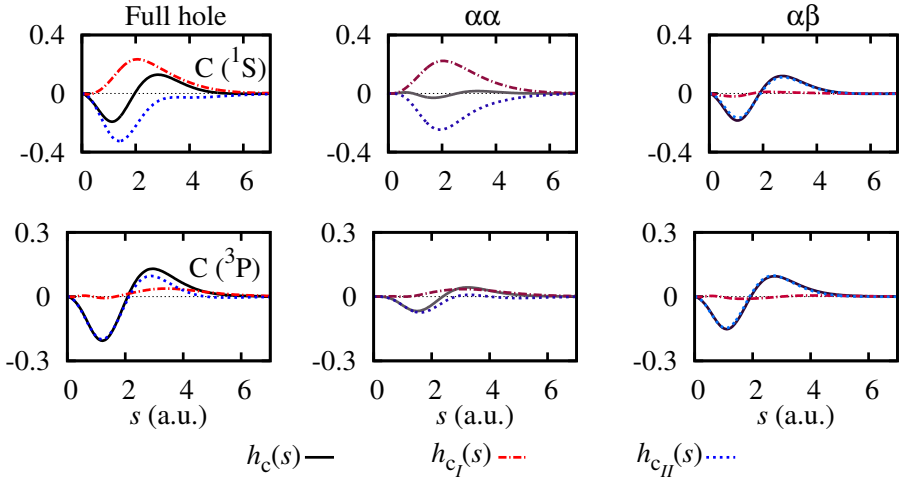


Figure 8: The Coulomb hole, $h_c(s)$, and its correlation components, $h_{c_I}(s)$ and $h_{c_{II}}(s)$, (left column), their same-spin contributions to the Coulomb hole (middle column), and opposite-spin contributions (right column) for singlet state carbon (top row) and the triplet, ground state (bottom row).

The reason behind this characteristic profiles is explored through the spin decomposition of the holes. We have used the carbon atom as a representative example of these three atoms. As we can see in Figure 8, the spin components of the Coulomb hole present similar profiles in singlet carbon ($C(^1S)$) and singlet H_2 (compare to Figure 4). The shape of the Coulomb hole is mostly given by the opposite-spin $h_{c_{II}}(s)$ component and the same-spin components are more important in the description of the singlet than in the triplet. The latter feature can be explained from the nature of $\rho_2^{\alpha\alpha,\mathrm{SD}}(\mathbf{1},\mathbf{2})$, which contributes to $h_{c_I}^{\alpha\alpha}(s)$. Since $h_{c_I}^{\alpha\alpha}(s)$ is greater than zero, its magnitude can be related to the inability of $\rho_2^{\alpha\alpha,\mathrm{SD}}(\mathbf{1},\mathbf{2})$ to reproduce the number of electron pairs of the same spin

upon integration over $\mathbf{1}$ and $\mathbf{2}$. The larger the electron correlation effects on the first-order density, the larger the deviation from the number of electron pairs of the same spin.[92] Therefore, singlet carbon, which is much more affected by nondynamic correlation, presents larger values of $h_{c_I}^{\alpha\alpha}(s)$ and, since the total Coulomb hole is mostly given by the opposite-spin component, $h_{c_{II}}^{\alpha\alpha}(s)$ will be large and (at least partially) compensate for $h_{c_I}^{\alpha\alpha}(s)$ (see Figure 8). In Table 1, it is shown that the trace deviation of the like-spin SD pair density is larger in the singlet than in the triplet.

Table 1: Same-spin pair-density traces for Ne, the singlet and triplet states of carbon atom. $\mathrm{Tr}\left[\rho_2^{\alpha\alpha}(\mathbf{1},\mathbf{2})\right] = N^\alpha(N^\alpha - 1)$.

	Ne	C(^{3}P)	C(^{1}S)
$\mathrm{Tr}\left[\rho_2^{\alpha\alpha}(\mathbf{1},\mathbf{2})\right]$	20.00	7.00	6.00
$\mathrm{Tr}\left[\rho_2^{\alpha\alpha,\mathrm{SD}}(\mathbf{1},\mathbf{2})\right]$	20.06	7.11	6.59

4.4 The Be_2 Coulomb hole

The description of both Be atom and its dimer is very sensitive to the level of theory and the quality of the basis set employed[93] (see also Supporting Information, where we follow a previous strategy[81,94] to analyze the effect of the basis set). Due to the well-known near-degeneracy of $2s$ and $2p$ orbitals in Be, the binding of both fragments and, in general, the potential energy curve (PEC) of Be_2 has been the subject of study of both experimentalists and computational chemists alike. One of the most intriguing features of the PEC of Be_2 is a change of slope that generates two potential minima, one corresponding to the equilibrium geometry at 4.72 a.u., and the second one at twice the equilibrium distance. Because of this double well, a single-reference method with a basis set that includes neither d nor f functions usually gives the second geometry as the equilibrium one, or even a repulsive, non-binding curve.[95–97] In our study, we have run a frozen-core full configuration interactions (FCI) calculation with the aug-cc-pVTZ basis set to study the Coulomb hole. It has been shown elsewhere[96] that the core electrons do not play an important role in the electronic description of Be_2 because the triple and quadruple excitations of the valence electrons are the ones

responsible for most of the electron correlation.

Figure 9 contains the Coulomb holes of two geometries of Be_2. Two interesting features of $h_c(s)$ are the presence of a small *bump* around the hole evaluated at the interatomic distance, $h_c(s = R)$, and a shoulder at the small s region. The latter is connected with lack of correlation of the core electrons (compare to Figure S2 and see Ref. 94 for a discussion of the effect of core electrons in the Coulomb hole). For both geometries, $h_{c_I}(s)$ takes large numbers due to the nondynamic correlation arisen from the near-degeneracy between the $2s$ and $2p$ orbitals of beryllium atom; this multireference character is preserved in the dimer. It has actually been reported that the valence molecular orbitals $3\sigma_g$, and $3\sigma_u$, which arise from linear combinations of $2s$ and $2p$, show significant orbital occupation numbers and, hence, should be considered in the calculation.[96,97] Indeed, the FCI occupation numbers of $3\sigma_g$ and $3\sigma_u$ are 0.0589 at $R = 24.57$ a.u.. Conversely, since the 2σ orbitals are completely occupied in the HF description, it cannot provide an accurate 1-RDM at this stretched geometry and, therefore, $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ is large. Certainly, the wrong dissociating description provided by HF, plus the strong multireference character of the dissociated Be fragments, makes the HF 1-RDM a very poor reference. The long-range part of the Coulomb hole of stretched Be_2 presents a maximum that belongs to $h_{c_I}(s)$, caused by the presence of the valence electrons localized near their respective nuclei. The long-range part of $h_{c_{II}}(s)$ in stretched Be_2 presents a minuscule yet positive area as a result of dispersion interactions.

The most remarkable feature of the Coulomb hole splitting of this molecule is the dominance of both types of correlation at short range. In most of the cases studied thus far, the short-range part of the Coulomb hole has always been dominated by $h_{c_{II}}(s)$, and it defined the shape of the short-range part of the total hole $h_c(s)$. Instead, the shape of the short-range part of the total hole in Be_2 at the equilibrium and stretched geometries is molded by both correlation contributions.

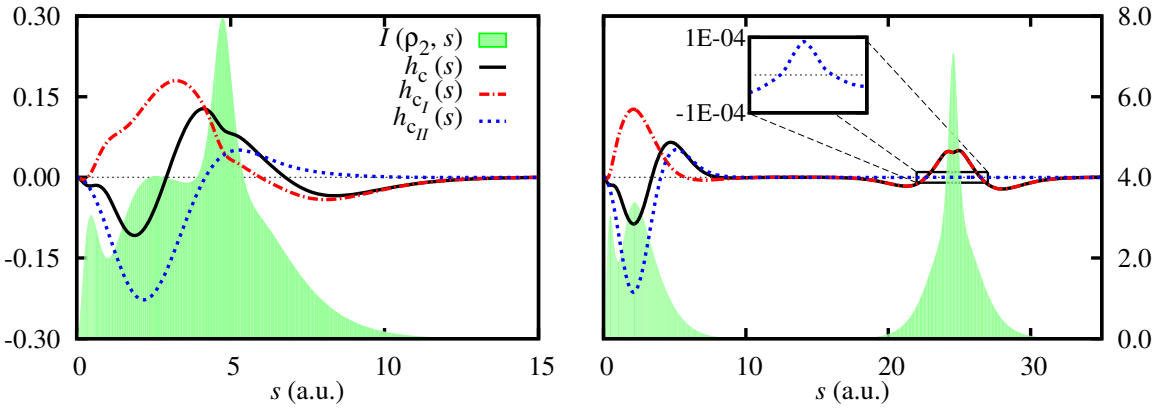


Figure 9: Coulomb hole (black), $h_{c_I}(s)$ (red) and $h_{c_{II}}(s)$ (blue) correlation components, and the intracule density (shadowed green region, right y-axis) of the Be_2 molecule at 4.72 a.u. (left) and 24.57 a.u. (right) bond lenghts.

4.5 Types A and B of nondynamic correlation: Be(Z) and H_2

Hollett and Gill recognized two types of nondynamic correlation that are classified according to the ability of the unrestricted Hartree–Fock (UHF) method to capture them.[11] The first type, type A, arises from the absolute near degeneracies, for example those that occur when a homolytic dissociation takes place, as in the stretching of a diatomic molecule. As the interatomic separation distance R increases, the energy gap ($\Delta\epsilon_{gap}$) between the highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO, respectively) becomes smaller until it turns zero and such orbitals become degenerate. Certainly, RHF cannot describe the localization of the electrons at each nucleus, whereas the unrestricted formalism succeeds in doing so by getting rid of the ionic description. The second type, named type B, arises in the presence of relative near-degeneracies. In the Be isoelectronic series, the HOMO-LUMO gap widens as the effective charge Z increases. However, the difference between gap increments, $\Delta\left(\Delta\epsilon_{gap}\right)$, remains constant. In this scenario, UHF is not able to portray the nondynamic correlation in such molecules and multireference methods are required for an accurate description. Identifying both correlation types, A and B, is a challenging test for correlation indicators.[20] In the present section, we study whether correlation types A and B can be detected by $h_{c_I}(s)$. Namely, we analyze a typical case of type B correlation, the isoelectronic series of Be, Be(Z) with Z = 3–8, and the dissociated and equilibrium geometries

of H_2 as an example of type A correlation.

In the literature, there exist many studies about the consistency and convergence of basis sets using the Be atom due to the difficulty to correctly reproduce short-range interactions in this system.[95,98] We have optimized an even-tempered basis set of 10 *s, p* and *d* functions to perform the calculations of Be(Z). Information about the optimization is provided in the Supporting Information.

It has been demonstrated elsewhere that the correlation energy in Be grows linearly with Z due to the $2s - 2p$ near-degeneracy.[99] Consequently, one would expect that the energy difference between UHF and RHF, ΔE, increases with Z. Instead, Hollet and Gill demonstrated the contrary in their work: the molecules present a triplet instability and UHF can describe nondynamic correlation from $Z = 3.0$ to $Z = 4.25$, yielding lower total energies with respect to RHF. Conversely, the fraction of energy recovered by UHF decreases as Z increases and the UHF description in $5 \leq Z \leq 8$ is actually equivalent to the RHF (see Table 2). The HOMO-LUMO gap widens with the effective charge Z, but the difference between gaps, $\Delta\left(\Delta\epsilon_{\text{gap}}\right)$, remains constant indicating a linear growth of the gap.

Table 2: The Restricted Hartree–Fock energy, Unrestricted Hartree–Fock energy, their energy difference, the spin contamination from the UHF calculation, the HOMO and LUMO energies and their difference, and the difference between energy gaps for Be-like ions with $3 \leq Z \leq 8$.

Z	E_{RHF}	E_{UHF}	ΔE	$\langle S^2 \rangle_{\text{UHF}}$	ϵ_{HOMO}	ϵ_{LUMO}	$\Delta\epsilon_{\text{gap}}$	$\Delta\left(\Delta\epsilon_{\text{gap}}\right)$
3	-7.38012	-7.39086	-0.01074	0.7132	0.04521	0.02639	0.21864	0.15874
4	-14.57022	-14.57059	-0.00036	0.1378	-0.31300	0.06438	0.37738	0.22385
5	-24.23339	-24.23339	0.00000	0.0000	-0.87412	-0.27289	0.60123	0.22636
6	-36.39601	-36.39601	0.00000	0.0000	-1.69410	-0.86651	0.82759	0.22368
7	-57.07217	-57.07217	0.00000	0.0000	-2.76685	-1.71558	1.05127	0.21915
8	-68.22939	-68.22939	0.00000	0.0000	-4.08835	-2.81793	1.27042	-

Figure 10 contains the $h_{c_I}(s)$ contribution for Be-like ions with $3 \leq Z \leq 6$. In accordance

to the facts summarized in Table 2 and in Ref. 11, the UHF and RHF holes for $Z = 5$ and 6 are indistinguishable, as the UHF calculation bears the restricted solution. Instead, the UHF and RHF holes for the Be atom are quite similar, and for $Z = 3$ both descriptions are no longer equivalent.

Figure S3 of the Supporting Information provides the Coulomb hole decomposition of a UHF calculation of the H_2 with a minimal basis set. The latter is barely indistinguishable from the FCI counterpart (Figure 2), indicating that $h_{c_I}(s)$ accounts also for type A correlation. In this sense, $h_{c_I}(s)$ can describe nondynamic correlation, regardless of its type.

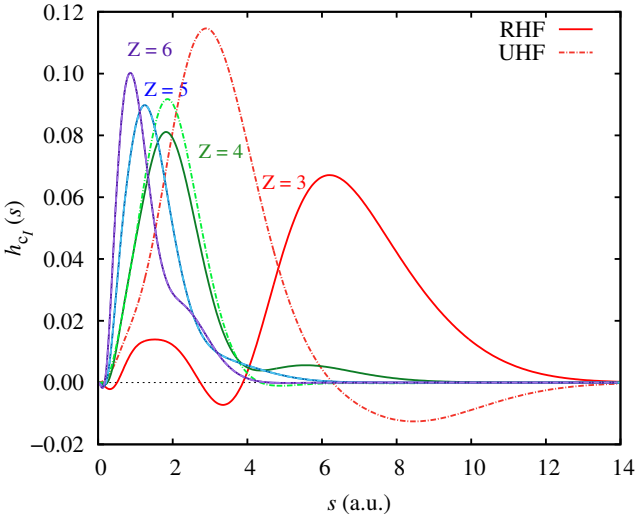


Figure 10: $h_{c_I}(s)$ component of the Coulomb hole of Be-like atoms, Be(Z), with $Z = 3$, 4, 5 and 6. Solid (dashed) lines depict the holes described using as reference the restricted (unrestricted) formalism of Hartree–Fock, that is, $h_{c_I}(s) = I\left(\rho_2^{\mathrm{SD}}, s\right) - I\left(\rho_2^{\mathrm{RHF}}, s\right)$ for solid lines, and $h_{c_I}(s) = I\left(\rho_2^{\mathrm{SD}}, s\right) - I\left(\rho_2^{\mathrm{UHF}}, s\right)$ for dashed lines.

4.6 C_{3v} H_4

We considered the PEC of H_4 in the C_{3v} symmetry point group, being the coordinate the distance between the axial H (H_{ax} hereinafter) and the equatorial plane where the rest of H are located, as it is illustrated in Figure 11.[100–106] In this study, we keep the distance between the equatorial hydrogens (labeled as R_{H_Δ} in Figure 11) fixed at 1.77 a.u.[104] and consider the lowest-lying states.

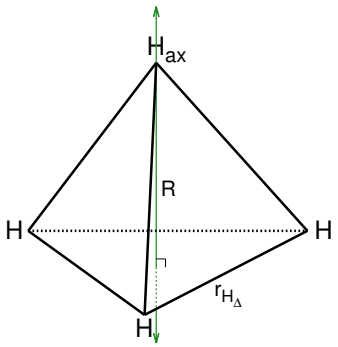


Figure 11: Graphical representation of H_4 at the C_{3v} symmetry point group, where the position of the axial hydrogen (H_{ax}) through the axial axis R (marked as green) represents the PEC analyzed. We keep r_{H_Δ} fixed at 1.77 a.u. bond distance. $R = 0$ when H_{ax} is at the equatorial plane.

The FCI potential energy curve (PEC) presents an intersystem crossing around $R = 4.50$ a.u. from S_0 to state S_1 (see Figure 12). At the S_0 state, the system is described as two interacting H_2 molecules, and the increase of R causes one of these molecules to dissociate, leaving H_3^+ in the equatorial plane and a H^-, H_{ax}, separated from it. The HF description of S_0 at the ground state is dominant in the CI vector, being the largest occupation numbers of the natural orbitals 1.957 and 1.936. Conversely, at $R = 28.35$ a.u. (S_1), the CI vector is composed by at least four highly contributing configurations, and the natural occupations are 1.954, 1.000, and 0.992, the 1.000 occupation corresponding to H_{ax}.

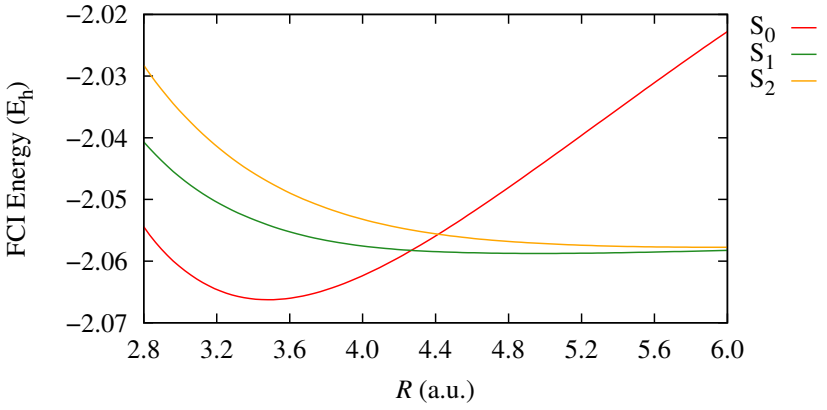


Figure 12: Potential energy curves (PEC) of H_4 at the C_{3v} symmetry point group, where the coordinate axis represents the separation of the axial hydrogen H_{ax} from the plane. A crossing from S_0 to S_1 takes place at 4.27 a.u.

The magnitudes of the Coulomb holes of equilibrium and dissociated geometries reflect the changes in the CI vector, where the S_1 state is more correlated than S_0. The correlation contributions depicted in Figure 13 show that $h_{c_{II}}(s)$ is prevalent at the equilibrium geometry and defines the shape of the Coulomb hole. At the stretched geometry, however, the short-range part of the Coulomb hole is dominated by $h_{c_I}(s)$, mainly caused by the multireference description of the S_1 state. $h_{c_{II}}(s)$ retains a simple shape and presents a peak at this geometry caused by the long-range dispersion interaction between H_{ax} and the rest of hydrogens in the plane, which is captured by $h_{c_{II}}(s)$. $h_{c_I}(s)$ at the stretched geometry presents two features we have not seen thus far in the first component of the Coulomb hole: it has a very large value at short-ranges and it shows important negative values at long range. These uncommon characteristics of the first component of the Coulomb hole put forward the very deficient description of HF, which does not dissociate into fragments with an integer number of electrons. If HF would dissociate into the correct number of electrons for each fragment, the shape of $h_{c_I}(s)$ would not show these features (compare to Figure S6). The large positive short-range part of $h_{c_I}(s)$ reflects the excess of electrons that HF locates in the H_3 plane. HF also locates less than one electron in H_{ax}, resulting in the negative long-range part of $h_{c_I}(s)$.

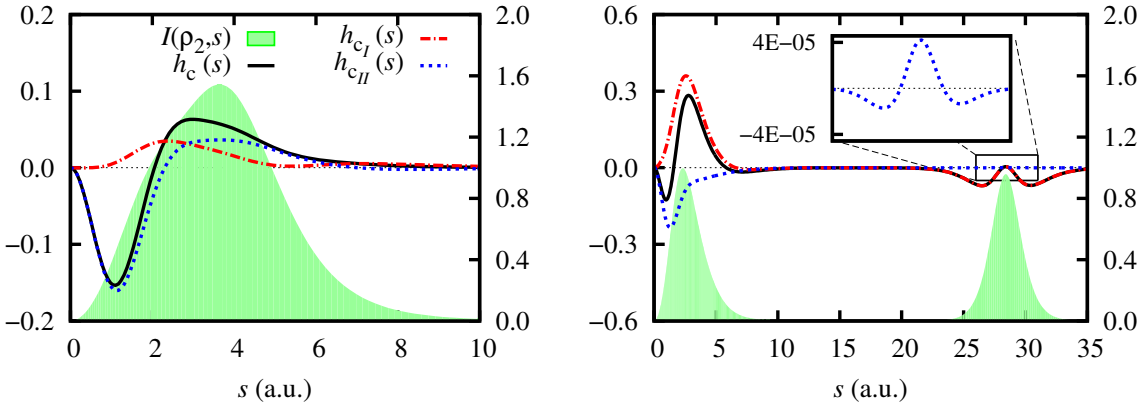


Figure 13: Coulomb hole (black), $h_{c_I}(s)$ (red) and $h_{c_{II}}(s)$ (blue) correlation components, and the intracule density (shadowed green region, right y-axis) of H_4 molecule at the C_{3v} symmetry point group with fixed r_{H_Δ}. H_{ax} is placed at (left) R = 3.50 a.u. (right) and R = 28.35 a.u. from the equatorial plane (see Figure 11 for details).

4.7 LiH and the harpoon mechanism

LiH presents an ionic ground state and its lowest-lying excited state is covalent.[107] Thus, in the adiabatic representation, the ground state of LiH presents an ionic bonding at equilibrium, with Li^+ and H^-, whereas the character of the state changes from ionic to covalent as the molecule dissociates (see Figure 14). The mechanism depicting the electron transition from hydrogen to lithium is known as the harpoon mechanism,[108] and it is caused by the small ionization potential of lithium and the large electron affinity of hydrogen. The potential energy curve presents an avoided crossing where the transition from the covalent state to the ionic description occurs, when Li and H are separated 6.8 a.u.[107]

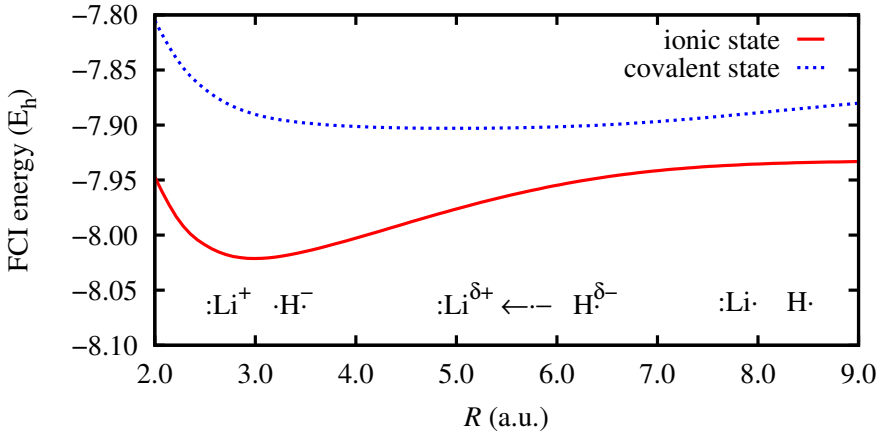


Figure 14: The potential energy curves of the lowest-lying diabatic states in LiH. An avoided crossing takes places around 6.8 a.u. Below, the harpoon mechanism taking place is depicted.

Two different profiles are obtained at different distances (see Figure 15) which, despite the different electronic distribution, resemble the C_{3v} H_4 profiles depicted in Figure 13. Whereas the equilibrium geometry is properly described by HF due to its ionic nature, the covalent character of the state at large interatomic distances is not. Therefore, the hole at equilibrium does not differ qualitatively from the hole of other molecules, such as H_2. On the other hand, at large distances, $h_{c_I}(s)$ is predominant along the interelectronic distance coordinate, with a modest area of $h_{c_{II}}(s)$ at short range and inappreciable at long range. Nevertheless, as we can see in the inset plot, the latter

is characterized by a maximum, which features the universal signature of dispersion. In this case, HF also dissociates the LiH molecule into a non-integer number of fragments and, therefore, the resemblance between this hole and the one of H_4. In fact, at the stretched geometry, the difference between the exact intracular and the UHF one is negligible (see Figure S5), indicating that all the features of the Coulomb hole in Figure 15 are entirely due to the wrong dissociation of HF.

[22,67]

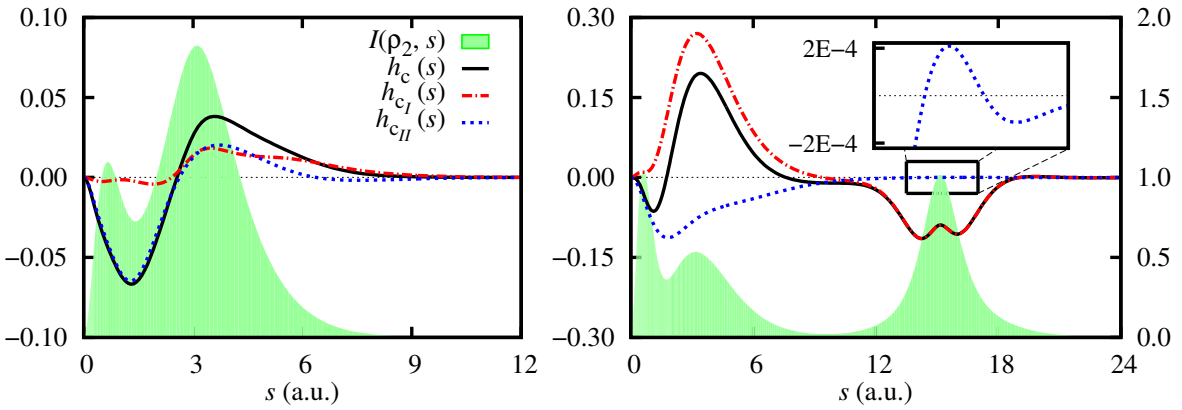


Figure 15: Coulomb hole (black), $h_{c_I}(s)$ (red) and $h_{c_{II}}(s)$ (blue) correlation components, and the intracule density (shadowed green region, right y-axis) of LiH at the equilibrium geometry at 3.02 a.u., and 15.12 a.u. bond length.

5 Conclusions

We have studied a correlation decomposition scheme that provides a natural separation of the pair density. Upon integration over the intracule coordinate, a range-separation of the Coulomb hole into two components, h_{c_I} and $h_{c_{II}}$, arises. The c_I component describes the long-range interactions that are nondynamic in nature, whereas the c_{II} component accounts for dynamic and nondynamic short-range interactions. These components are exhaustively analyzed throughout various examples that put forward the most relevant features of this approach and how they can be used to develop efficient ways to capture electron correlation.

First of all, through the asymptotic properties of the hole parts, we explain how the component based only on the first-order reduced density matrix can retrieve the long-range part of electron correlation. Second, we perform an exhaustive analysis of the hydrogen molecule in a minimal basis set, dissecting the hole contributions into spin components. Third, we analyze the simplest molecule presenting a dispersion interaction, triplet H_2, and how $h_{c_{II}}$ helps to identify it. This dispersion signature is also identified in all the other molecules studied in this work, highlighting its universal character.

We also analyze the Coulomb holes of several atoms in different spin states, finding that the Coulomb hole components permit to distinguish clear correlation regimes that are not apparent from the whole Coulomb hole. Indeed, atoms with different spin states present the same Coulomb hole profile but their correlation components unravel relevant differences. The same-spin elements are the ones that capture nondynamic correlation in the system, and give away the true multireference character of the singlet state of carbon and nitrogen atoms.

Finally, we analyze the two types of nondynamic correlation, types A and B,[11] and show that they can be both captured by the first component of the Coulomb hole. Profiles of $h_{c_I}(s)$ calculated with an unrestricted reference clearly differentiate from profiles calculated with a restricted wavefunction when type B nondynamic correlation is present. The larger the presence of type B nondynamic correlation, the bigger the difference between both holes. Interestingly, the correlation indicators based in natural orbitals that were also developed from this partition[20] could not make such distinction.

The results of this work hold the promise to aid in the development of new electronic structure methods that efficiently capture electron correlation. In particular, the models of the c_{II} component can be combined with other results available in the literature[56,66,109] to develop new approximations in the reduced density matrix functional theory,[53,54] although the present work is not limited to this

theory.

Acknowledgement

The authors thank Prof. Manuel Yáñez and Prof. Paul W. Ayers for proposing the study of Be_2 and H_4 at the C_{3v} symmetry point group, respectively. This research was funded by Spanish MINECO grant numbers PGC2018-098212-B-C21, EUIN2017-88605 and EUR2019-103825, the Basque Government / Eusko Jaurlaritza (GV/EJ) grant number IT1254-19, PIBA19-0004, and 2019-CIEN-000092-01. M. V. N. acknowledges the doctoral grant BES-2015-072734 by Spanish MINECO and E.R.C. acknowledges funding from the Juan de la Cierva program IJCI-2017-34658.

Supporting Information Available

The Supporting Information contains the Coulomb hole of the He-Ne atomic series, a study of the basis set consistency for the Be atom, the short-range region of the Coulomb hole of the beryllium molecule, a spin decomposition of the Coulomb hole of H_2 in a minimal basis set, and a more detailed development of the asymptotics and conditions of $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ for the hydrogen molecule in a minimal basis.

References

(1) Löwdin, P.-O. 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. *Phys. Rev.* **1955**, *97*, 1474–1489.

(2) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advances Electronic Structure Theory*; Courier Corporation: New York, 2012.

(3) Wilson, S. *Electron correlation in molecules*; Dover: New York, 2007.

(4) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. Explicitly correlated electrons in molecules. *Chem. Rev.* **2012**, *112*, 4–74.

(5) Tew, D. P.; Klopper, W.; Helgaker, T. Electron correlation: The many-body problem at the heart of chemistry. *J. Comput. Chem.* **2007**, *28*, 1307–1320.

(6) Cremer, D. Density functional theory: coverage of dynamic and non-dynamic electron correlation effects. *Mol. Phys.* **2001**, *99*, 1899–1940.

(7) Ziesche, P. On relations between correlation, fluctuation and localization. *J. Mol. Struct. (Theochem)* **2000**, *527*, 35–50.

(8) Raeber, A.; Mazziotti, D. A. Large eigenvalue of the cumulant part of the two-electron reduced density matrix as a measure of off-diagonal long-range order. *Phys. Rev. A* **2015**, *92*, 052502.

(9) Benavides-Riveros, C. L.; Lathiotakis, N. N.; Schilling, C.; Marques, M. A. Relating correlation measures: The importance of the energy gap. *Phys. Rev. A* **2017**, *95*, 032507.

(10) Gottlieb, A. D.; Mauser, N. J. New measure of electron correlation. *Phys. Rev. Lett.* **2005**, *95*, 123003.

(11) Hollett, J. W.; Gill, P. M. The two faces of static correlation. *J. Chem. Phys.* **2011**, *134*, 114111.

(12) Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. A simple DFT-based diagnostic for nondynamical correlation. *Theor. Chem. Acc.* **2013**, *132*, 1–9.

(13) Proynov, E.; Liu, F.; Kong, J. Analyzing effects of strong electron correlation within Kohn-Sham density-functional theory. *Phys. Rev. A* **2013**, *88*, 032510.

(14) Cioslowski, J. Density-driven self-consistent-field method: Density-constrained correlation energies in the helium series. *Phys. Rev. A* **1991**, *43*, 1223–1228.

(15) Valderrama, E.; Ludeña, E. V.; Hinze, J. Analysis of dynamical and nondynamical components of electron correlation energy by means of local-scaling density-functional theory. *J. Chem. Phys.* **1997**, *106*, 9227–9235.

(16) Valderrama, E.; Ludeña, E. V.; Hinze, J. Assessment of dynamical and nondynamical correlation energy components for the beryllium-atom isoelectronic sequence. *J. Chem. Phys.* **1999**, *110*, 2343–2353.

(17) Mok, D. K. W.; Neumann, R.; Handy, N. C. Dynamic and Nondynamic Correlation. *J. Phys. Chem.* **1996**, *100*, 6225–6230.

(18) Benavides-Riveros, C. L.; Lathiotakis, N. N.; Marques, M. A. Towards a formal definition of static and dynamic electronic correlations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 12655–12664.

(19) Vuckovic, S.; Irons, T. J. P.; Wagner, L. O.; Teale, A. M.; Gori-Giorgi, P. Interpolated energy densities, correlation indicators and lower bounds from approximations to the strong coupling limit of DFT. *Phys. Chem. Chem. Phys.* **2017**, *19*, 6169–6183.

(20) Ramos-Cordoba, E.; Salvador, P.; Matito, E. Separation of dynamic and nondynamic correlation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24015–24023.

(21) Ramos-Cordoba, E.; Matito, E. Local Descriptors of dynamic and nondynamic correlation. *J. Chem. Theory Comput.* **2017**, *13*, 2705–2711.

(22) Via-Nadal, M.; Rodríguez Mayorga, M.; Ramos-Cordoba, E.; Matito, E. Singling Out Dynamic and Nondynamic Correlation. *J. Phys. Chem. Lett.* **2019**, *10*, 4032–4037.

(23) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.

(24) Sinanoğlu, O. Many-Electron Theory of Atoms, Molecules and Their Interactions. *Adv. Chem. Phys.* **1964**, *6*, 315–412.

(25) Stone, A. J. *The theory of intermolecular forces*, 2nd ed.; Oxford University Press, 2013.

(26) Shavitt, I. *Methods of electronic structure theory*; Springer, 1977; pp 189–275.

(27) Čížek, J. On the correlation problem in atomic and molecular systems. Calculation of wave-function components in Ursell-type expansion using quantum-field theoretical methods. *J. Phys. Chem.* **1966**, *45*, 4256–4266.

(28) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.

(29) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(30) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd: Chichester, 2000.

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

(32) White, S. R.; Huse, D. A. Numerical renormalization-group study of low-lying eigenstates of the antiferromagnetic S=1 Heisenberg chain. *Phys. Rev. B* **1993**, *48*, 3844.

(33) Frenkel, J. *Wave Mechanics, Volume 2: Advanced General Theory*; The Clarendon Press: Oxford, 1934; pp 460–462.

(34) Roos, B. O.; Siegbahn, P. In *Methods of electronic structure theory*; Schaefer III, F., Ed.; Plenum Press: New York, 1977.

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

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

(37) Buenker, R. J.; Peyerimhoff, S. D. Individualized configuration selection in CI calculations with subsequent energy extrapolation. *Theor. Chim. Acta (Berlin)* **1974**, *35*, 33–58.

(38) Werner, H.-J.; Knowles, P. J. An efficient internally contracted multiconfiguration–reference configuration interaction method. *J. Chem. Phys.* **1988**, *89*, 5803–5814.

(39) Pastorczak, E.; Hapka, M.; Veis, L.; Pernal, K. Capturing the Dynamic Correlation for Arbitrary Spin-Symmetry CASSCF Reference with Adiabatic Connection Approaches: Insights into the Electronic Structure of the Tetramethyleneethane Diradical. *J. Phys. Chem. Lett.* **2019**, *10*, 4668–4674.

(40) Hollett, J. W.; Loos, P.-F. Capturing static and dynamic correlation with Δ NO-MP2 and Δ NO-CCSD. *J. Chem. Phys.* **2020**, *152*, 014101.

(41) Savin, A. In *Recent Developments of Modern Density Functional Theory*; Seminario, J. M., Ed.; Elsevier: Amsterdam, 1996; p 327.

(42) Savin, A. A combined density functional and configuration interaction method. *Int. J. Quantum Chem.* **1988**, *34*, 59–69.

(43) Fromager, E.; Toulouse, J.; Jensen, H. J. A. On the universality of the long-/short-range separation in multiconfigurational density-functional theory. *J. Chem. Phys.* **2007**, *126*, 074111.

(44) Toulouse, J.; Gerber, I. C.; Jansen, G.; Savin, A.; Angyán, J. G. Adiabatic-Connection Fluctuation-Dissipation Density-Functional Theory based on range separation. *Phys. Rev. Lett.* **2009**, *102*, 096404.

(45) Toulouse, J.; Colonna, F.; Savin, A. Long-range short-range separation of the electron-electron interaction in density-functional theory. *Phys. Rev. A* **2004**, *70*, 062505.

(46) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

(47) Baer, R.; Livshits, E.; Salzner, U. Tuned range-separated hybrids in density functional theory. *Ann. Rev. Phys. Chem.* **2010**, *61*, 85–109.

(48) Garrett, K.; Sosa Vazquez, X.; Egri, S. B.; Wilmer, J.; Johnson, L. E.; Robinson, B. H.; Isborn, C. M. Optimum exchange for calculation of excitation energies and hyperpolarizabilities of organic electro-optic chromophores. *J. Chem. Theory Comput.* **2014**, *10*, 3821–3831.

(49) Bao, J. J.; Gagliardi, L.; Truhlar, D. G. Multiconfiguration pair-density functional theory for doublet excitation energies and excited state geometries: the excited states of CN. *Phys. Chem. Chem. Phys.* **2017**, *19*, 30089–30096.

(50) Gagliardi, L.; Truhlar, D. G.; Li Manni, G.; Carlson, R. K.; Hoyer, C. E.; Bao, J. L. Multiconfiguration pair-density functional theory: A new way to treat strongly correlated systems. *Acc. Chem. Res.* **2017**, *50*, 66–73.

(51) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration pair-density functional theory. *J. Chem. Theory Comput.* **2014**, *10*, 3669–3680.

(52) Coleman, A. J. Density matrices in the quantum theory of matter: Energy, intracules and extracules. *Int. J. Quantum Chem.* **1967**, *1*, 457–464.

(53) Piris, M.; Ugalde, J. M. Perspective on natural orbital functional theory. *Int. J. Quantum Chem.* **2014**, *114*, 1169–1175.

(54) Pernal, K.; Giesbertz, K. J. H. Reduced Density Matrix Functional Theory (RDMFT) and Linear Response Time-Dependent RDMFT (TD-RDMFT). *Top. Curr. Chem.* **2015**, *368*, 125–183.

(55) Ramos-Cordoba, E.; Lopez, X.; Piris, M.; Matito, E. H_4: A challenging system for natural orbital functional approximations. *J. Chem. Phys.* **2015**, *143*, 164112.

(56) Ramos-Cordoba, E.; Salvador, P.; Piris, M.; Matito, E. Two new constraints for the cumulant matrix. *J. Chem. Phys.* **2014**, *141*, 234101.

(57) McWeeny, R. *Methods of Molecular Quantum Mechanics*, 2nd ed.; Academic Press: London, 1989.

(58) Thakkar, A. J.; Smith Jr, V. H. Form factors and total scattering intensities for the helium-like ions from explicitly correlated wavefunctions. *J. Phys. B: At. Mol. Opt. Phys.* **1978**, *11*, 3803.

(59) Thakkar, A. J.; Tripathi, A.; Smith, V. H. Anisotropic electronic intracule densities for diatomics. *Int. J. Quantum Chem.* **1984**, *26*, 157–166.

(60) Thakkar, A. J.; Tripathi, A.; Smith Jr, V. H. Molecular X-ray and electron-scattering intensities. *Phys. Rev. A* **1984**, *29*, 1108.

(61) Valderrama, E.; Fradera, X.; Ugalde, J. M. Determination of the integrated X-ray scattering intensities through the electron-pair relative-motion density at the origin. *Phys. Rev. A* **2001**, *64*, 044501.

(62) Coulson, C. A. Present state of molecular structure calculations. *Rev. Mod. Phys.* **1960**, *32*, 170–177.

(63) Coulson, C. A.; Neilson, A. H. Electron correlation in the ground state of helium. *Proc. Phys. Soc. London* **1961**, *78*, 831–837.

(64) Boyd, R. F.; Coulson, C. A. Coulomb hole in some excited states of helium. *J. Phys. B: At. Mol. Opt. Phys.* **1973**, *6*, 782–793.

(65) Valderrama, E.; Ugalde, J.; Boyd, R. Many-Electron Densities and Reduced Density Matrices. 2000.

(66) Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Via-Nadal, M.; Piris, M.; Matito, E. Comprehensive benchmarking of density matrix functional approximations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 24029–24041.

(67) Via-Nadal, M.; Rodríguez-Mayorga, M.; Matito, E. A Salient Signature of van der Waals Interactions. *Phys. Rev. A* **2017**, *96*, 050501.

(68) Coleman, A. J.; Yukalov, V. I. *Reduced density matrices: Coulson's challenge*; Springer Verlag: Berlin, 2000; Vol. 72.

(69) Coleman, A. J. Structure of Fermion Density Matrices. *Rev. Mod. Phys.* **1963**, *35*, 668–687.

(70) Mazziotti, D. A. Structure of fermionic density matrices: complete *N*-representability conditions. *Phys. Rev. Lett.* **2012**, *108*, 263002.

(71) Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Feixas, F.; Matito, E. Electron correlation effects in third-order densities. *Phys. Chem. Chem. Phys.* **2017**, *19*, 4522–4529.

(72) Ayers, P. W.; Davidson, E. R. Linear inequalities for diagonal elements of density matrices. *Adv. Chem. Phys.* **2007**, *134*, 443–483.

(73) Feixas, F.; Solà, M.; Barroso, J. M.; Ugalde, J. M.; Matito, E. New Approximation to the Third-Order Density. Application to the Calculation of Correlated Multicenter Indices. *J. Chem. Theory Comput.* **2014**, *10*, 3055–3065.

(74) Lieb, E. H. Variational principle for many-fermion systems. *Phys. Rev. Lett.* **1981**, *46*, 457.

(75) Levy, M. *Density Matrices and Density Functionals*; Springer, 1987; pp 479–498.

(76) Buijse, M. A. Thesis: Electron Correlation. Fermi and Coulomb holes, dynamical and nondynamical correlation. Ph.D. thesis, Vrije Universiteit, Amsterdam, The Netherlands, 1991.

(77) Kutzelnigg, W.; Mukherjee, D. Cumulant expansion of the reduced density matrices. *J. Chem. Phys.* **1999**, *110*, 2800–2809.

(78) Mazziotti, D. A. Approximate solution for electron correlation through the use of Schwinger probes. *Chem. Phys. Lett.* **1998**, *289*, 419–427.

(79) March, N. H.; Pucci, R. Asymptotic form of first-order density matrix for atoms and molecules. *J. Chem. Phys.* **1981**, *75*, 496–497.

(80) Knowles, P. J.; Handy, N. C. A determinant based full configuration interaction program. *Comput. Phys. Commun.* **1989**, *54*, 75.

(81) Matito, E.; Cioslowski, J.; Vyboishchikov, S. F. Properties of harmonium atoms from FCI calculations: Calibration and benchmarks for the ground state of the two-electron species. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6712–6716.

(82) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. Electron affinities of the first -row atoms revisited. Systematic basis sets and wavefunctions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

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

(84) Matito, E.; Feixas, F. DMn program. 2009; Universitat de Girona (Spain) and Uniwersytet Szczeciński (Poland).

(85) Rodríguez-Mayorga, M. RHO2-OPS: 2-DM Operations. 2016; Institut de Química Computacional i Catàlisi (IQCC), Universitat of Girona, Catalonia, Spain.

(86) Cioslowski, J.; Liu, G. Fast evaluation of electron intracule and extracule densities on large grids of points. *J. Chem. Phys.* **1996**, *105*, 4151–4158.

(87) Hehre, W. J.; Stewart, R. F.; Pople, J. A. . *J. Chem. Phys.* **1969**, *51*, 2657.

(88) Matito, E.; Duran, M.; Solà, M. A Novel Exploration of the Hartree–Fock Homolytic Bond Dissociation Problem in the Hydrogen Molecule by Means of Electron Localization Measures. *J. Chem. Educ.* **2006**, *83*, 1243.

(89) Gritsenko, O.; Baerends, E. J. A simple natural orbital mechanism of "pure" van der Waals interaction in the lowest excited triplet state of the hydrogen molecule. *J. Chem. Phys.* **2006**, *124*, 054115.

(90) Bowman Jr, J. D.; Hirschfelder, J. O.; Wahl, A. C. Extended Hartree–Fock Calculations for the Ground State and Hartree–Fock Calculations for the First Excited State of H_2. *J. Chem. Phys.* **1970**, *53*, 2743–2749.

(91) Janowski, T.; Pulay, P. A Benchmark Comparison of σ/σ and π/π Dispersion: the Dimers of Naphthalene and Decalin, and Coronene and Perhydrocoronene. *J. Am. Chem. Soc.* **2012**, *134*, 17520–17525.

(92) Matito, E.; Solà, M.; Salvador, P.; Duran, M. Electron sharing indexes at the correlated level. Application to aromaticity calculations. *Faraday Discuss.* **2007**, *135*, 325–345.

(93) NIST Computational Chemistry Comparison and Benchmark Database,
NIST Standard Reference Database Number 101
Release 20, August 2019, Editor: Russell D. Johnson III.
http://cccbdb.nist.gov/,.

(94) Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Lopez, X.; Solà, M.; Ugalde, J. M.; Matito, E. The Coulomb Hole of the Ne Atom. *ChemistryOpen* **2019**, *8*, 411–417.

(95) Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. Electronic structure analysis of the ground-state potential energy curve of Be2. *J. Phys. Chem. A* **2010**, *114*, 8687–8696.

(96) El Khatib, M.; Bendazzoli, G. L.; Evangelisti, S.; Helal, W.; Leininger, T.; Tenti, L.; Angeli, C. Beryllium dimer: A bond based on non-dynamical correlation. *J. Phys. Chem. A* **2014**, *118*, 6664–6673.

(97) Noga, J.; Kutzelnigg, W.; Klopper, W. CC-R12, a correlation cusp corrected coupled-cluster

method with a pilot application to the Be_2 potential curve. *Chem. Phys. Lett.* **1992**, *199*, 497–504.

(98) Gálvez, F.; Buendía, E.; Sarsa, A. Two-electron properties for the beryllium atom from explicitly correlated wavefunctions. *Chem. Phys. Lett.* **2003**, *378*, 330–336.

(99) Chakravorty, S. J.; Gwaltney, S. R.; Davidson, E. R.; Parpia, F. A.; p Fischer, C. F. Ground-state correlation energies for atomic ions with 3 to 18 electrons. *Phys. Rev. A* **1993**, *47*, 3649.

(100) Boothroyd, A.; Martin, P.; Keogh, W.; Peterson, M. An accurate analytic H_4 potential energy surface. *J. Chem. Phys.* **2002**, *116*, 666–689.

(101) Aguado, A.; Suárez, C.; Paniagua, M. Accurate global fit of the H_4 potential energy surface. *J. Chem. Phys.* **1994**, *101*, 4004–4010.

(102) Montgomery Jr, J. A.; Michels, H. H. On the stability of H_4 in C_{3v} symmetry. *J. Chem. Phys.* **1987**, *86*, 5882–5883.

(103) Evleth, E. M.; Kassab, E. On the stability of H_4 in C_{3v} symmetry: Comments. *J. Chem. Phys.* **1988**, *89*, 3928–3929.

(104) Jorge, F. E. Ab Initio Valence Bond Calculations of the Ground and First Two Singlet Excited States of H_4 Molecule. *Journal of the Brazilian Chemical Society* **1993**, *4*, 26–29.

(105) Theodorakopoulos, G.; Petsalakis, I. D.; Nicolaides, C. A. Potential energy hypersurfaces of H_4 in the ground and the first two singlet excited electronic states. *J. Mol. Struct. (Theochem)* **1987**, *149*, 23–31.

(106) Nicolaides, C. A.; Theodorakopoulos, G.; Petsalakis, I. D. Theory of chemical reactions of vibronically excited H_2 ($B^1\Sigma_u^+$). I. Prediction of a strongly bound excited state of H_4. *J. Chem. Phys.* **1984**, *80*, 1705–1706.

(107) Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Salvador, P.; Solà, M.; Matito, E. Bonding description of the Harpoon Mechanism. *Mol. Phys.* **2016**, *114*, 1345–1355.

(108) Polanyi, J. C.; Zewail, A. H. Direct observation of the transition state. *Acc. Chem. Res.* **1995**, *28*, 119–132.

(109) Cioslowski, J.; Piris, M.; Matito, E. Robust validation of approximate 1-matrix functionals with few-electron harmonium atoms. *J. Chem. Phys.* **2015**, *143*, 214101.

Supporting Information: Range separation of the Coulomb hole

Mireia Via-Nadal,[†,‡] Mauricio Rodríguez-Mayorga,[†] Eloy Ramos-Cordoba,[*,†,‡] and Eduard Matito[*,†,¶]

†*Donostia International Physics Center (DIPC), 20018 Donostia, Euskadi, Spain*

‡*Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Euskadi, Spain*

¶*IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Euskadi, Spain*

E-mail: eloy.raco@gmail.com; ematito@gmail.com

1 Basis set consistency in full configuration interactions calculations of Be atom

In previous studies of the Be atom and its dimer,[1–3] it has been found that the description of the species is very sensitive to the basis set consistency and it affects, particularly, to the short-range description. A systematic improvement of the basis set does not guarantee a better wavefunction (see, for instance, the ground state geometries for different methods and basis sets in the NIST database).[4] In this study, we have optimized an even-tempered basis set of *n*s, *n*p and *n*d functions to perform the calculations of Be-like ions, where *n* represents the number of orbitals used of each kind (these basis sets are referred as *n*SPD hereafter).[5] Figure S1 displays the Coulomb hole $h_c(s)$ of the beryllium atom calculated with 6SPD, 7SPD, 8SPD, and 10SPD basis sets (9SPD is not shown due to its high resemblance to 10SPD). $h_c(s)$ with Dunning's augmented basis functions, aug-cc-pVDZ and aug-cc-pVTZ, and Pople's 6-311G** have also been calculated for comparison.

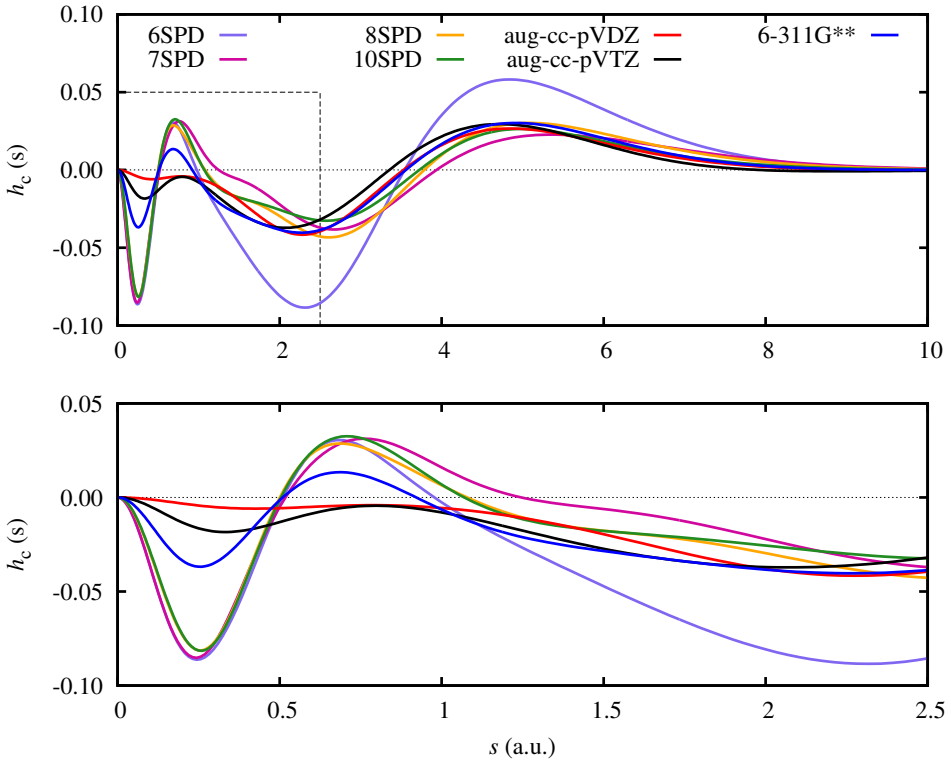


Figure S1: Coulomb hole $h_c(s)$ of Be computed with different basis sets. The graph below depicts a zoomed section from the plot above, marked with dashed lines.

Note that, as the quality of the basis set increases (from aug-cc-pVDZ to aug-cc-pVTZ), the short range part of $h_c(s)$ shows a sinusoidal shape that all even-tempered basis sets already reproduce (see the bottom plot in Figure S1). In a previous study, our group found that 6-311G** offers a poor description of the short-range part of the Coulomb hole, and that optimized basis sets provided a better description for core electrons.[3] Therefore, we have chosen the 10SPD basis set to study the rest of Be-like ions as the 9SPD one was indistinguishable from 10SPD (*i.e.* the basis set is converged with respect to the number of *s*, *p* and *d* basis functions).

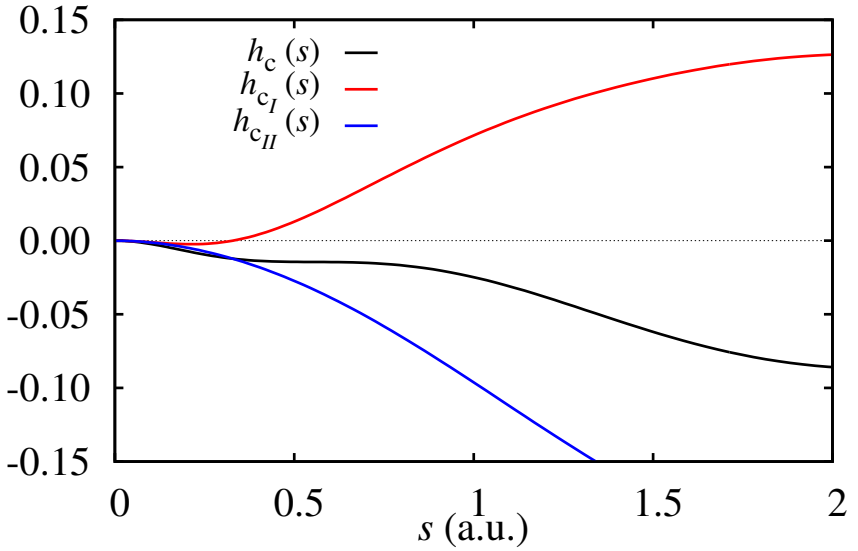


Figure S2: The short-range region of the Coulomb hole ($h_c(s)$), $h_{c_I}(s)$, and $h_{c_{II}}(s)$ of Be_2 at 24.57 a.u. of bond distance.

2 The short-range region of the Coulomb hole of Be_2

The short-range region of the Coulomb hole of the beryllium dimer does not present a smooth decrease as the rest of Coulomb holes exposed in the article. Figure S2 contains this short-range region for a stretched geometry of Be_2, where the Coulomb hole $h_c(s)$ has an inflection point instead of a smooth minimum, which is usually defined as a shoulder-shape profile. The correlation decomposition of the Coulomb hole indicates that $h_{c_I}(s)$ is zero in the short-range region, with a posterior even growth, and $h_{c_{II}}(s)$ decreases smoothly with the interelectronic distance s. The combination of both components results in a shoulder-like shape in the short-range region of h_c.

3 Type A of nondynamic correlation Coulomb hole of stretched H_2

Because UHF has the ability to describe type A nondynamic correlation, molecular dissociations such as the one of H_2 are correctly accounted for by this method. There is no type A nondynamic

correlation in the equilibrium geometry of H_2, where we have not yet reached the Coulson-Fisher point. On the other hand, the bond stretching gives rise to type A nondynamic correlation and, eventually, the UHF intracule in a minimal basis set is exactly coincident with the FCI one. Thus, the Coulomb holes are identical, as depicted in Figure S3.

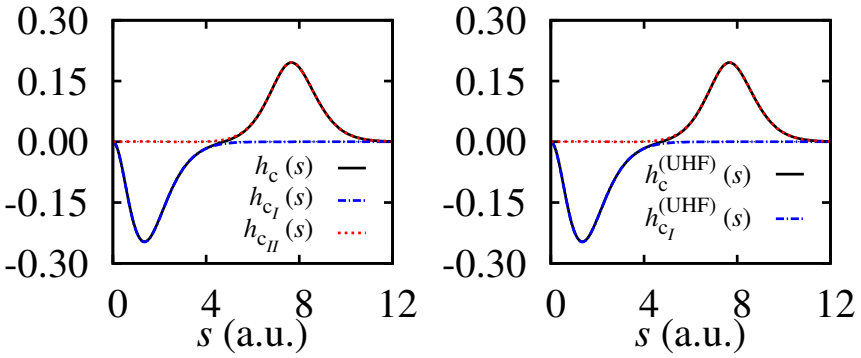


Figure S3: (Left) Coulomb hole $h_c(s)$, $h_{c_I}(s)$, and $h_{c_{II}}(s)$ of H_2 at $R = 7.56$ a.u. in a minimal basis set, as presented in Figure 2 of the manuscript. (Right) Coulomb hole calculated using the UHF intracule pair density instead of the FCI one: $h_c^{(\mathrm{UHF})}(s) = I(\rho_2^{\mathrm{UHF}}, s) - I(\rho_2^{\mathrm{RHF}}, s)$, and $h_{c_{II}}^{(\mathrm{UHF})}(s) = I(\rho_2^{\mathrm{UHF}}, s) - I(\rho_2^{\mathrm{SD}}, s)$.

4 Coulomb hole of He-Ne Atomic series

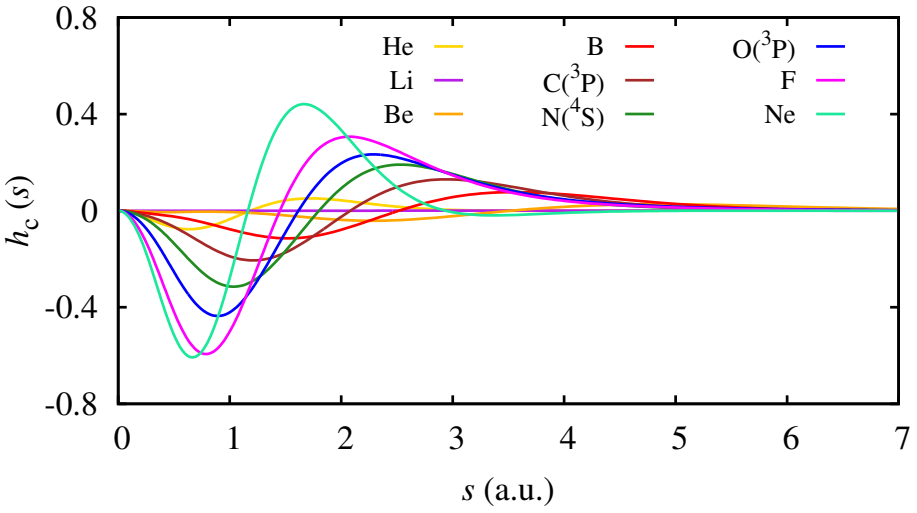


Figure S4: Coulomb hole $h_c(s)$ of the He-Ne atomic series.

Coulomb holes $h_c(s)$ for the He-Ne atomic series in their ground state are presented in Figure S4. Notice that the dimension of lithium's Coulomb hole is smaller compared to the rest of atomic series and shows as a flat line.

5 Cases with RHF dissociating to an incorrect number of electrons per fragment.

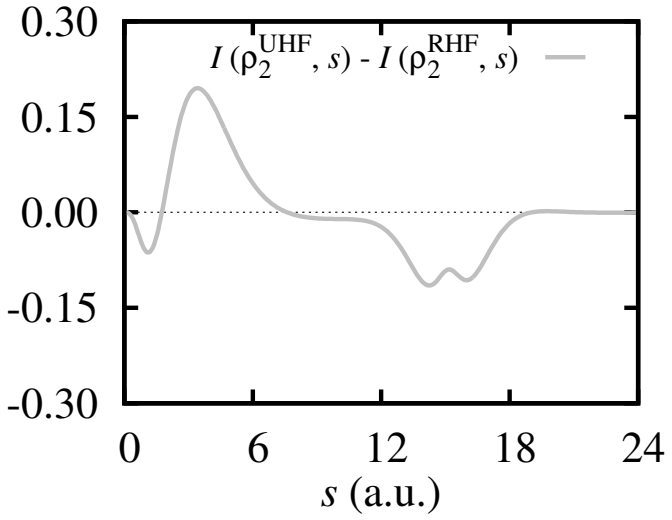


Figure S5: The intracule probability density difference of the UHF and RHF pair densities, $I(\rho_2^{\mathrm{UHF}}, s) - I(\rho_2^{\mathrm{RHF}}, s)$, of the LiH molecule at $R = 15.12$ a.u..

Because RHF is not able to properly dissociate heterogeneous diatomic molecules with different electronegativities, the dissociated fragments of LiH are not neutral; namely, RHF dissociates the molecule into ions, Li+ and H-. Because of this, the RHF intracule probability density unusually presents an excess of electron-pair probability at the long-ranges of interelectronic distance caused by the extra electron in H. Instead, the unrestricted formalism correctly separates LiH into two neutral atoms. The UHF wavefunction (and also the intracule probability density) provides a close description of the FCI one. Figure S5 contains the intracule difference between the UHF and

RHF. This difference is indistinguishable from the Coulomb hole h_c presented in the manuscript, due to the quality of the UHF description.

A similar situation occurs in the trigonal planar H_4 molecule (C_{3v} symmetr point group). The HF description describes the axial hydrogen with a non-integer number of electrons. The UHF wavefunction provides a better description of the molecule, and because of this the magnitude of the corresponding hole is smaller. Note that, however, H_4 is a more correlated molecule than LiH and, therefore, the unrestricted formalism of HF is not able to provide a correct description to the system either.

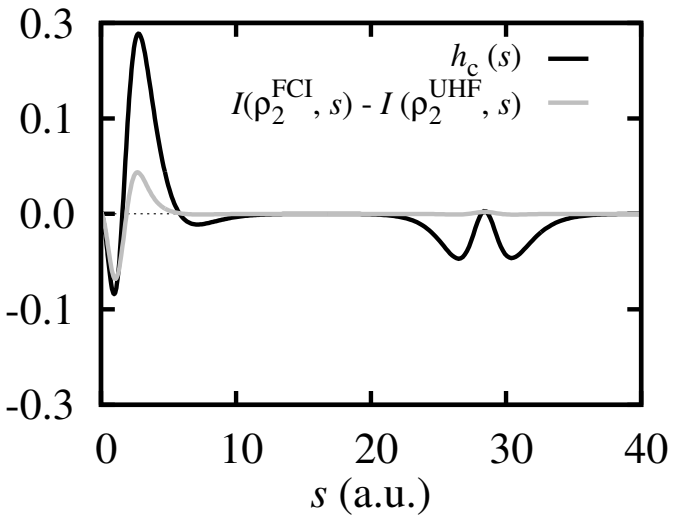


Figure S6: Coulomb hole h_c (black) and the intracule difference between the FCI and the UHF pair densities $I(\rho_2^{\mathrm{FCI}}, s) - I(\rho_2^{\mathrm{UHF}}, s)$ (gray) of H_4 molecule at C_{3v} symmetry point group, with $_{\mathrm{H}_\Delta}$ = 1.77 a.u., and H_{ax} placed at R = 28.35 a.u. from the equatorial plane (see also Figure 11 of the paper).

6 Conditions for $\Delta\rho_2^{\mathbf{c}_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{\mathbf{c}_{II}}(\mathbf{1},\mathbf{2})$: H_2 in a minimal basis

Let us analyze $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ for the hydrogen dimer described in a minimal basis set with a normalized $1s$ orbital centered on each nuclei A and B. The bonding and antibonding

orthonormal orbitals that arise are

$$\phi_1(\boldsymbol{r}) = \frac{1}{\sqrt{2(1+S_{AB})}}(S_A(\boldsymbol{r}) + S_B(\boldsymbol{r})) \tag{S1}$$

and

$$\phi_2(\boldsymbol{r}) = \frac{1}{\sqrt{2(1-S_{AB})}}(S_A(\boldsymbol{r}) - S_B(\boldsymbol{r})), \tag{S2}$$

respectively, and are eigenfunctions of the Fock operator. S_{AB} is the overlap between both atomic orbitals.

The Slater determinant built with the lowest-lying molecular orbital, eq S1, corresponds to the (normalized) Hartree-Fock ground-state, $\Phi_0(\mathbf{1},\mathbf{2})$:

$$\Phi_0(\mathbf{1},\mathbf{2}) = |1\overline{1}\rangle = \frac{1}{\sqrt{2}}\left(\phi_1(\mathbf{1})\overline{\phi_1}(\mathbf{2}) + \overline{\phi_1}(\mathbf{1})\phi_1(\mathbf{2})\right), \tag{S3}$$

where the bar over the atomic orbital indicates that the spin state of the electron is beta (otherwise, it is alpha). The doubly-excited Slater determinant $\Phi_{\mathrm{D}}(\mathbf{1},\mathbf{2})$ reads

$$\Phi_{\mathrm{D}}(\mathbf{1},\mathbf{2}) = |2\overline{2}\rangle = \frac{1}{\sqrt{2}}\left(\phi_2(\mathbf{1})\overline{\phi_2}(\mathbf{2}) + \overline{\phi_2}(\mathbf{1})\phi_2(\mathbf{2})\right), \tag{S4}$$

and with a linear combination of both determinants one can construct the exact wavefunction for H_2 in a minimal basis set:

$$\Psi(\mathbf{1},\mathbf{2}) = c_0\Phi_0(\mathbf{1},\mathbf{2}) + c_{\mathrm{D}}\Phi_{\mathrm{D}}(\mathbf{1},\mathbf{2}), \tag{S5}$$

and, to keep the wavefunction normalized we impose $c_0^2 + c_{\mathrm{D}}^2 = 1$.

From this wavefunction, one can obtain its second-order reduced density matrix (2-RDM),

$$\begin{aligned}
\rho_2(\mathbf{1},\mathbf{2};\mathbf{1'},\mathbf{2'}) &= 2\Psi^*(\mathbf{1},\mathbf{2})\,\Psi(\mathbf{1'},\mathbf{2'}) \\
&= c_0^2\Big(\phi_1^*(\mathbf{1})\overline{\phi_1^*}(\mathbf{2})\phi_1(\mathbf{1'})\overline{\phi_1}(\mathbf{2'}) + \overline{\phi_1^*}(\mathbf{1})\phi_1^*(\mathbf{2})\overline{\phi_1}(\mathbf{1'})\phi_1(\mathbf{2'}) \\
&\quad + \phi_1^*(\mathbf{1})\overline{\phi_1^*}(\mathbf{2})\overline{\phi_1}(\mathbf{1'})\phi_1(\mathbf{2'}) + \overline{\phi_1^*}(\mathbf{1})\phi_1^*(\mathbf{2})\phi_1(\mathbf{1'})\overline{\phi_1}(\mathbf{2'})\Big) \\
&\quad + c_{\mathrm{D}}^2\Big(\phi_2^*(\mathbf{1})\overline{\phi_2^*}(\mathbf{2})\phi_2(\mathbf{1'})\overline{\phi_2}(\mathbf{2'}) + \overline{\phi_2^*}(\mathbf{1})\phi_2^*(\mathbf{2})\overline{\phi_2}(\mathbf{1'})\phi_2(\mathbf{2'}) \\
&\quad + \phi_2^*(\mathbf{1})\overline{\phi_2^*}(\mathbf{2})\overline{\phi_2}(\mathbf{1'})\phi_2(\mathbf{2'}) + \overline{\phi_2^*}(\mathbf{1})\phi_2^*(\mathbf{2})\phi_2(\mathbf{1'})\overline{\phi_2}(\mathbf{2'})\Big) \\
&\quad + c_0 c_{\mathrm{D}}\Big(\phi_1^*(\mathbf{1})\overline{\phi_1^*}(\mathbf{2})\phi_2(\mathbf{1'})\overline{\phi_2}(\mathbf{2'}) + \overline{\phi_1^*}(\mathbf{1})\phi_1^*(\mathbf{2})\overline{\phi_2}(\mathbf{1'})\phi_2(\mathbf{2'}) \\
&\quad + \phi_1^*(\mathbf{1})\overline{\phi_1^*}(\mathbf{2})\overline{\phi_2}(\mathbf{1'})\phi_2(\mathbf{2'}) + \overline{\phi_1^*}(\mathbf{1})\phi_1^*(\mathbf{2})\phi_2(\mathbf{1'})\overline{\phi_2}(\mathbf{2'}) \\
&\quad + \phi_2^*(\mathbf{1})\overline{\phi_2^*}(\mathbf{2})\phi_1(\mathbf{1'})\overline{\phi_1}(\mathbf{2'}) + \overline{\phi_2^*}(\mathbf{1})\phi_2^*(\mathbf{2})\overline{\phi_1}(\mathbf{1'})\phi_1(\mathbf{2'}) \\
&\quad + \phi_2^*(\mathbf{1})\overline{\phi_2^*}(\mathbf{2})\overline{\phi_1}(\mathbf{1'})\phi_1(\mathbf{2'}) + \overline{\phi_2^*}(\mathbf{1})\phi_2^*(\mathbf{2})\phi_1(\mathbf{1'})\overline{\phi_1}(\mathbf{2'})\Big);
\end{aligned} \tag{S6}$$

the first-order reduced density matrix (1-RDM), which is obtained upon integration of the 2-RDM over the second-electron coordinate,

$$\begin{aligned}
\rho_1(\mathbf{1};\mathbf{1'}) &= \int d\mathbf{2}\,\rho_2(\mathbf{1},\mathbf{2};\mathbf{1'},\mathbf{2'})\,|_{\mathbf{2'}=\mathbf{2}} \\
&= c_0^2\left(\phi_1^*(\mathbf{1})\phi_1(\mathbf{1'}) + \overline{\phi_1^*}(\mathbf{1})\overline{\phi_1}(\mathbf{1'})\right) + c_{\mathrm{D}}^2\left(\phi_2^*(\mathbf{1})\phi_2(\mathbf{1'}) + \overline{\phi_2^*}(\mathbf{1})\overline{\phi_2}(\mathbf{1'})\right);
\end{aligned} \tag{S7}$$

and the electron density, which is the diagonal of the 1-RDM,

$$\rho(\mathbf{1}) = \rho_1(\mathbf{1};\mathbf{1'}) = c_0^2\left(|\phi_1(\mathbf{1})|^2 + |\overline{\phi_1}(\mathbf{1})|^2\right) + c_{\mathrm{D}}^2\left(|\phi_2(\mathbf{1})|^2 + |\overline{\phi_2}(\mathbf{1})|^2\right). \tag{S8}$$

The HF pair density is

$$\rho_2^{\mathrm{HF}}(\mathbf{1},\mathbf{2}) = 2\Phi_0^*(\mathbf{1},\mathbf{2})\,\Phi_0(\mathbf{1},\mathbf{2}) = |\phi_1(\mathbf{1})|^2|\overline{\phi_1}(\mathbf{2})|^2 + |\overline{\phi_1}(\mathbf{1})|^2|\phi_1(\mathbf{2})|^2 \tag{S9}$$

and the single-determinant approximation, according to eq 6 in the manuscript,[6] reads

$$\begin{aligned}
\rho_2^{\mathrm{SD}}(\mathbf{1},\mathbf{2}) =& \rho(\mathbf{1})\rho(\mathbf{2}) - |\rho_1(\mathbf{1};\mathbf{2})|^2 \\
=& c_0^4\left(|\phi_1(\mathbf{1})|^2|\overline{\phi_1}(\mathbf{2})|^2 + |\overline{\phi_1}(\mathbf{1})|^2|\phi_1(\mathbf{2})|^2\right) + c_{\mathrm{D}}^4\left(|\phi_2(\mathbf{1})|^2|\overline{\phi_2}(\mathbf{2})|^2 + |\overline{\phi_2}(\mathbf{1})|^2|\phi_2(\mathbf{2})|^2\right) \\
&+c_0^2c_{\mathrm{D}}^2\Big(|\phi_1(\mathbf{1})|^2|\overline{\phi_2}(\mathbf{2})|^2 + |\overline{\phi_1}(\mathbf{1})|^2|\phi_2(\mathbf{2})|^2 + |\phi_1(\mathbf{1})|^2|\phi_2(\mathbf{2})|^2 + |\overline{\phi_1}(\mathbf{1})|^2|\overline{\phi_2}(\mathbf{2})|^2 \\
&+|\phi_1(\mathbf{2})|^2|\overline{\phi_2}(\mathbf{1})|^2 + |\overline{\phi_1}(\mathbf{2})|^2|\phi_2(\mathbf{1})|^2 + |\phi_1(\mathbf{2})|^2|\phi_2(\mathbf{1})|^2 + |\overline{\phi_1}(\mathbf{2})|^2|\overline{\phi_2}(\mathbf{1})|^2 \\
&-\phi_1^*(\mathbf{1})\phi_1(\mathbf{2})\phi_2^*(\mathbf{1})\phi_2(\mathbf{2}) - \overline{\phi_1^*}(\mathbf{1})\overline{\phi_1}(\mathbf{2})\overline{\phi_2^*}(\mathbf{1})\overline{\phi_2}(\mathbf{2}) \\
&-\phi_1^*(\mathbf{2})\phi_1(\mathbf{1})\phi_2^*(\mathbf{2})\phi_2(\mathbf{1}) - \overline{\phi_1^*}(\mathbf{2})\overline{\phi_1}(\mathbf{1})\overline{\phi_2^*}(\mathbf{2})\overline{\phi_2}(\mathbf{1})\Big),
\end{aligned} \tag{S10}$$

At the particular case when the two centers are far apart (and electrons are located at each center), the overlap between the atomic orbitals becomes zero, *i.e.* $|R_{AB}| \to \infty \Rightarrow S_{AB} \to 0$. Then, it is legitimate to consider that $\phi_1(A) \approx \phi_2(A)$ and $-\phi_1(B) \approx \phi_2(B)$, A and B standing for the position of the corresponding atom. Therefore, eq S10 becomes

$$\begin{aligned}
\lim_{|R_{AB}|\to\infty} \rho_2^{\mathrm{SD}}(A,B) =& \left(c_0^4 + c_{\mathrm{D}}^4 + 2c_0^2c_{\mathrm{D}}^2\right)\left(|\phi_1(A)|^2|\overline{\phi_1}(B)|^2 + |\overline{\phi_1}(A)|^2|\phi_1(B)|^2\right) \\
&+ 2c_0^2c_{\mathrm{D}}^2\Big(|\phi_1(A)|^2|\phi_1(B)|^2 + |\overline{\phi_1}(A)|^2|\overline{\phi_1}(B)|^2 \\
&+ \phi_1^*(A)\phi_1(B)\phi_1^*(A)\phi_1(B) + \overline{\phi_1^*}(A)\overline{\phi_1}(B)\overline{\phi_1^*}(A)\overline{\phi_1}(B)\Big) \\
=& \rho_2^{\mathrm{HF}}(A,B) + 4c_0^2c_{\mathrm{D}}^2\Big(|\phi_1(A)|^2|\phi_1(B)|^2 + |\overline{\phi_1}(A)|^2|\overline{\phi_1}(B)|^2\Big),
\end{aligned} \tag{S11}$$

where, for the sake of simplicity, we have considered real orbitals, $\phi_i^*(\boldsymbol{r}) = \phi_i(\boldsymbol{r})$. Hence,

$$\lim_{|R_{AB}|\to\infty} \Delta\rho_2^{\mathrm{c}_I}(A,B) = \rho_2^{\mathrm{SD}}(A,B) - \rho_2^{\mathrm{HF}}(A,B) = 4c_0^2c_{\mathrm{D}}^2\Big(|\phi_1(A)|^2|\phi_1(B)|^2 + |\overline{\phi_1}(A)|^2|\overline{\phi_1}(B)|^2\Big). \tag{S12}$$

When the wavefunction is only composed of the HF determinant, $c_{\mathrm{D}} = 0 \Rightarrow \Delta\rho_2^{\mathrm{c}_I}(A,B) \to 0$. Conversely, at the limit of the dissociation, the expansion coefficients are equal in weight, $c_0 = -c_{\mathrm{D}} = \frac{1}{\sqrt{2}}$. Hence, the spin-integrated c_I component of the pair density tends to a value which is as large as the HF pair density, $\Delta\rho_2^{\mathrm{c}_I}(A,B) \to \rho_2^{\mathrm{HF}}(A,B)$, which results in $\Delta\rho_2^{\mathrm{c}_I}(A,B)$ capturing the

nondynamic correlation arisen from the stretching.

One could also analyze the spin-components of these expressions. The hydrogen molecule only presents opposite-spin interactions, with just one alpha and just one beta electron in each nucleus. Thus the HF pair density does not have exchange terms and will read

$$\rho_2^{\mathrm{HF}}(\mathbf{1},\mathbf{2}) = \rho^{\mathrm{HF},\alpha}(\mathbf{1})\rho^{\mathrm{HF},\beta}(\mathbf{2}) + \rho^{\mathrm{HF},\beta}(\mathbf{1})\rho^{\mathrm{HF},\alpha}(\mathbf{2}). \tag{S13}$$

However, ρ_2^{SD} includes some same-spin elements by construction, and hence

$$\rho_2^{\mathrm{SD}}(\mathbf{1},\mathbf{2}) = \rho^{\alpha}(\mathbf{1})\rho^{\alpha}(\mathbf{2}) + \rho^{\beta}(\mathbf{1})\rho^{\beta}(\mathbf{2}) + \rho^{\alpha}(\mathbf{1})\rho^{\beta}(\mathbf{2}) + \rho^{\beta}(\mathbf{1})\rho^{\alpha}(\mathbf{2}) - |\rho_1^{\alpha\alpha}(\mathbf{1};\mathbf{2})|^2 - |\rho_1^{\beta\beta}(\mathbf{1};\mathbf{2})|^2. \tag{S14}$$

If we assume $\rho^{\mathrm{HF},\sigma}(\mathbf{1}) \approx \rho^{\sigma}(\mathbf{1})$, only the same-spin terms survive in the c_I component,

$$\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2}) \approx \rho^{\alpha}(\mathbf{1})\rho^{\alpha}(\mathbf{2}) + \rho^{\beta}(\mathbf{1})\rho^{\beta}(\mathbf{2}) - |\rho_1^{\alpha\alpha}(\mathbf{1};\mathbf{2})|^2 - |\rho_1^{\beta\beta}(\mathbf{1};\mathbf{2})|^2. \tag{S15}$$

When a dissociation occurs, namely when $R_{AB} \to \infty$, the $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ vanishes at short range (the first non-vanishing term of the short-range expansion is quadratic in r_{12}).

The asymptotics of the 1-RDM define the asymptotics of the c_I component of the pair density. March and Pucci[7] found that, when electrons are separated infinitely from each other and are also separated from any nucleus A, $\rho_1(\mathbf{1};\mathbf{2}) \to \sqrt{\rho(\mathbf{1})\rho(\mathbf{2})}$.[1] Therefore, the same-spin component of the SD approximation reduces to zero, and thus

$$\lim_{\substack{r_{12}\to\infty \\ r_{A1},r_{B2}\to\infty}} \Delta\rho_2^{c_I,\alpha\alpha}(\mathbf{1},\mathbf{2}) = 0 \qquad\qquad \forall A,B \tag{S16}$$

The opposite-spin component does not depend on the 1-RDM, therefore its asymptotics is defined

[1] For non-degenerate $(N-1)$-particle systems.[8]

by the product of the correlated and uncorrelated electron densities:

$$\lim_{\substack{r_{12}\to\infty \\ r_{A1},r_{B2}\to\infty}} \Delta\rho_2^{c_I,\alpha\beta}(\mathbf{1},\mathbf{2}) = \rho^{\alpha}(\mathbf{1})\rho^{\beta}(\mathbf{2}) - \rho^{\mathrm{HF},\alpha}(\mathbf{1})\rho^{\mathrm{HF},\beta}(\mathbf{2}) = 0, \qquad \forall A, B \tag{S17}$$

because the density dies off quickly far from the nuclei.

When electrons are placed on top of each other $r_{12} \to 0$ (the coalescence point), the same-spin component of $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ is zero by construction. Nevertheless, there exists a small probability of opposite-spin electrons to be on top of each other:

$$\lim_{r_{12}\to 0} \Delta\rho_2^{c_I,\alpha\beta}(\mathbf{1},\mathbf{2}) = \rho^{\alpha}(\mathbf{1})\rho^{\beta}(\mathbf{1}) - \rho^{\mathrm{HF},\alpha}(\mathbf{1})\rho^{\mathrm{HF},\beta}(\mathbf{1}). \tag{S18}$$

Now, let us consider the limits of $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$. The value of $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ at the coalescence point depends exclusively on the opposite-spin component, because the same-spin one vanishes due to the Pauli principle,

$$\lim_{r_{12}\to 0} \Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2}) = \Delta\rho_2^{c_{II},\alpha\beta}(\mathbf{1},\mathbf{1}) + \Delta\rho_2^{c_{II},\beta\alpha}(\mathbf{1},\mathbf{1}) = 2\left(\rho_2^{\alpha\beta}(\mathbf{1},\mathbf{1}) - \rho^{\alpha}(\mathbf{1})\rho^{\beta}(\mathbf{1})\right), \tag{S19}$$

and, hence, its behavior depends on the on-top pair density value.[9]

The actual pair density (the diagonal elements of eq S6) as $R_{AB} \to \infty$ is dominated by the value at the points close to the nuclei. Let us consider, $\rho_2(A,B)$,

$$\begin{aligned}\rho_2(A,B) =& c_0^2\left(|\phi_1(A)|^2|\overline{\phi_1}(B)|^2 + |\overline{\phi_1}(A)|^2|\phi_1(B)|^2\right) + c_{\mathrm{D}}^2\left(|\phi_1(A)|^2|\overline{\phi_1}(B)|^2 + |\overline{\phi_1}(A)|^2|\phi_1(B)|^2\right) \\ & -2c_0c_{\mathrm{D}}\left(|\phi_1(A)|^2|\overline{\phi_1}(B)|^2 + |\overline{\phi_1}(A)|^2|\phi_1(B)|^2\right) \\ =& (c_0 - c_{\mathrm{D}})^2\rho_2^{\mathrm{HF}}(A,B),\end{aligned} \tag{S20}$$

and therefore c_{II} component of the pair density is (see eq S11)

$$\Delta\rho_2^{c_{II}}(A,B) = \rho_2(A,B) - \rho_2^{\mathrm{SD}}(A,B) = (c_0 - c_{\mathrm{D}})^2\,\rho_2^{\mathrm{HF}}(A,B) - (1 + 4c_0^2c_{\mathrm{D}}^2)\rho_2^{\mathrm{HF}}(A,B), \tag{S21}$$

where we have assumed a closed-shell system.

At the dissociation limit $c_0 = -c_{\mathrm{D}} = \frac{1}{\sqrt{2}}$ and, therefore, $\lim_{R_{AB}\to\infty}\rho_2(A,B) = 2\rho_2^{\mathrm{HF}}(A,B)$. Hence, $\lim_{R_{AB}\to\infty}\Delta\rho_2^{c_{II}}(A,B) = 0$. Unlike $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$, the long-range component of $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ vanishes and, therefore, we have a convenient range separation of the pair density into c_I and c_{II} components.

References

(1) Gálvez, F.; Buendía, E.; Sarsa, A. Two-electron properties for the beryllium atom from explicitly correlated wavefunctions. *Chem. Phys. Lett.* **2003**, *378*, 330–336.

(2) Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. Electronic structure analysis of the ground-state potential energy curve of Be2. *J. Phys. Chem. A* **2010**, *114*, 8687–8696.

(3) Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Lopez, X.; Solà, M.; Ugalde, J. M.; Matito, E. The Coulomb Hole of the Ne Atom. *ChemistryOpen* **2019**, *8*, 411–417.

(4) NIST Computational Chemistry Comparison and Benchmark Database,
NIST Standard Reference Database Number 101
Release 20, August 2019, Editor: Russell D. Johnson III.
http://cccbdb.nist.gov/,.

(5) Matito, E.; Cioslowski, J.; Vyboishchikov, S. F. Properties of harmonium atoms from FCI calculations: Calibration and benchmarks for the ground state of the two-electron species. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6712–6716.

(6) Löwdin, P.-O. 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. *Phys. Rev.* **1955**, *97*, 1474–1489.

(7) March, N. H.; Pucci, R. Asymptotic form of first-order density matrix for atoms and molecules. *J. Chem. Phys.* **1981**, *75*, 496–497.

(8) Ernzerhof, M.; Burke, K.; Perdew, J. P. Long-range asymptotic behavior of ground-state wave functions, one-matrices, and pair densities. *J. Chem. Phys.* **1996**, *105*, 2798–2803.

(9) Perdew, J. P.; Ernzerhof, M.; Burke, K.; Savin, A. On-top pair-density interpretation of spin density functional theory, with applications to magnetism. *Int. J. Quantum Chem.* **1997**, *61*, 197–205.

Chapter 4

A new footprint of London dispersion interactions

PHYSICAL REVIEW A **96**, 050501(R) (2017)

Salient signature of van der Waals interactions

Mireia Via-Nadal,[1] Mauricio Rodríguez-Mayorga,[1,2] and Eduard Matito[1,3,*]

[1]*Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), and Donostia International Physics Center (DIPC), P.K. 1072, 20080 Donostia, Euskadi, Spain*

[2]*Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, University of Girona, 17071 Girona, Catalonia, Spain*

[3]*IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Euskadi, Spain*

(Received 11 October 2017; published 13 November 2017)

van der Waals interactions govern the physics of a plethora of molecular structures. It is well known that the leading term in the distance-based London expansion of the van der Waals energy for atomic and molecular dimers decays as $1/R^6$, where R is the dimer distance. Using perturbation theory, we find the leading term in the distance-based expansion of the intracule pair density at the interatomic distance. Our results unveil a universal $1/R^3$ decay, which is less prone to numerical errors than the $1/R^6$ dependency, and it is confirmed numerically in H_2 and He_2 molecules. This *signature* of van der Waals interactions can be directly used in the construction of approximate pair density and energy functionals including vdW corrections.

DOI: 10.1103/PhysRevA.96.050501

I. INTRODUCTION

Dispersion or van der Waals (vdW) interactions are ubiquitous in nature, governing the stability of molecules and materials [1] and having an essential role in molecular recognition [2], the double-helical structure of DNA [3], molecular adsorption on surfaces [4,5], and the adhesion of micromachined surfaces [6]. They are so important in physics, chemistry, and biology, that even the most simple electronic structure methods consider corrections for vdW interactions. Due to their long-range dynamic-correlation nature, they are not well modeled by standard functionals in density functional theory (DFT) [1], which by construction are essentially local or semilocal in nature [7]. Hence, except for a few functionals [8,9], most DFT functionals include *ad hoc* energy corrections to account for vdW interactions [10–12].

vdW forces arise from the electrostatic interaction between fluctuations in the electron density, and the pairwise effect in the energy shows a leading $1/R^6$ dependency, where R is the interaction distance between fragments. This fact is often exploited in the construction of effective pairwise potentials that enter the expressions of various methods. On the other hand, the effect of vdW interactions in the wave function or related quantities has been less discussed in the literature [13–16]. This knowledge could shed some light in the design of computational approaches including vdW interactions and provide further tests to calibrate electronic structure theory methods.

In this Rapid Communication we use perturbation theory to find the leading term in the expansion of the intracule pair density in terms of R, the interatomic distance. Our results reveal a universal $1/R^3$ dependency that is corroborated by numerical calculations in H_2 and He_2 molecules. Upon integration of the vdW contribution to the intracule we recover the vdW energy that follows the established $1/R^6$ dependency.

*Author to whom the correspondence should be addressed: ematito@gmail.com

II. THEORY

We start from the unperturbed wave function for a system of two hydrogen atoms, A and B, separated a distance R, given by the product of two hydrogenoid $1s$ functions [17],

$$\Psi^{(0)}(\mathbf{r}_1,\mathbf{r}_2) = N\phi_A(\mathbf{r}_1)\phi_B(\mathbf{r}_2). \tag{1}$$

For convenience we have chosen to represent the hydrogenoid functions by one Gaussian function of exponent α. This change does not alter the $1/R^6$ dependency on the calculation of the second-order correction to the energy (*vide infra*).

The perturbation operator contains all the interactions between fragments A and B:

$$\hat{H}^{(1)} = -\frac{1}{r_{A2}} - \frac{1}{r_{B1}} + \frac{1}{R} + \frac{1}{r_{12}}, \tag{2}$$

where we have labeled the electrons in A and B as 1 and 2, respectively (see Fig. 1). After application of Unsöld's approximation [18,19], the expansion of Eq. (2) in terms of $1/R$ at large distances yields the first-order correction to the wave function,

$$\Psi^{(1)}(\mathbf{r}_1,\mathbf{r}_2) \approx \frac{\Psi^{(0)}(\mathbf{r}_1,\mathbf{r}_2)}{R^3}(x_{A1}x_{B2} + y_{A1}y_{B2} - 2z_{A1}z_{B2}) + O(R^{-4}), \tag{3}$$

where we have assumed that the molecule is located in the z axis and x_{A1}, y_{B2}, etc. refer to the Cartesian components of the vectors in Eq. (2). Using the first-order correction to the wave function we can demonstrate that the leading term in the second-order correction to the energy follows the widely known $1/R^6$ dependency [17]

$$E^{(2)} = \frac{6}{\alpha^4 R^6} + O(R^{-8}). \tag{4}$$

Now, let us consider the first-order correction to the pair density,

$$n_2^{(1)}(\mathbf{r}_1,\mathbf{r}_2) = 2\Psi^{(0)}(\mathbf{r}_1,\mathbf{r}_2)\Psi^{(1)}(\mathbf{r}_1,\mathbf{r}_2). \tag{5}$$

 ©2017 American Physical Society

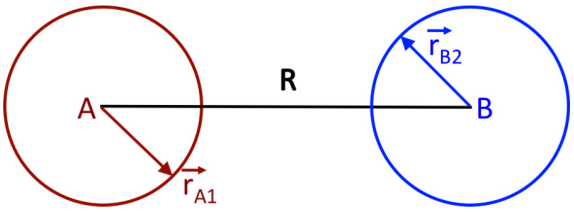


FIG. 1. Two atoms, A and B, separated by a distance R and two electrons with coordinates $\vec{r}_{A1}$ and $\vec{r}_{B2}$ defined with respect to the position of atoms A and B.

The intracule of the pair density is a function that provides a distribution of the interelectronic separations [20]

$$I(u) = \int\int d\mathbf{r}_1 d\mathbf{r}_2\, n_2(\mathbf{r}_1,\mathbf{r}_2)\delta(u - r_{12}), \tag{6}$$

and thus returns the average electron-electron distance upon integration over u. The intracule of the pair density is actually connected with an experimental observable, the x-ray scattering intensity, which is essentially determined by the Fourier-Bessel transform of the radial intracule probability density [21,22].

After some algebraic manipulation one can prove that the zeroth-order intracule of the pair density at R,

$$I^{(0)}(R) = \left(\frac{\alpha}{16\pi^3}\right)^{1/2}(1 - e^{-4\alpha R^2}), \tag{7}$$

yields a constant value in the limit, which corresponds to the distribution of two independent electrons. A Gaussian function enters the expression in Eq. (7) and, therefore, the form of the zeroth-order intracule function at R depends on the reference wave function, Eq. (1). Conversely, the first-order correction at R decays as $1/R^3$,

$$\lim_{R\to\infty} I^{(1)}(R) = -\frac{4(1+8\sqrt{2})\alpha^{5/2}}{\pi^{7/2}R^3}, \tag{8}$$

without any exponential terms in the expression, suggesting that the $1/R^3$ dependency does not rely on a particular choice of the zeroth-order reference. As we will check numerically in the next section, Eq. (8) puts forward a universal condition that can be employed to assess the performance of approximate pair densities and models of the intracule function in reproducing vdW interactions.

III. NUMERICAL EXAMPLES

As a zeroth-order pair density we choose

$$n_2^{\mathrm{SD}}(\mathbf{r}_1,\mathbf{r}_2) = n(\mathbf{r}_1)n(\mathbf{r}_2) - n_1(\mathbf{r}_1;\mathbf{r}_2)n_1(\mathbf{r}_2;\mathbf{r}_1), \tag{9}$$

where $n_1(\mathbf{r}_1;\mathbf{r}_2)$ is the first-order reduced density matrix. n_2^{SD} is the minimal model that guarantees that $n_2^{\mathrm{SD}}(\mathbf{r}_1,\mathbf{r}_2) \to n(\mathbf{r}_1)n(\mathbf{r}_2)$ at large interatomic distances and, at the same time, preserves the antisymmetric nature of particles [23], which is necessary to remove the spin entanglement effects that also appear at large interelectronic separations and are not included by the second-order perturbational treatment. Hence we will evaluate the intracule resulting from the following pair density difference

$$I^{(1)}(u) \approx \int\int d\mathbf{r}_1 d\mathbf{r}_2 \left[n_2(\mathbf{r}_1,\mathbf{r}_2) - n_2^{\mathrm{SD}}(\mathbf{r}_1,\mathbf{r}_2)\right]\delta(u - r_{12}). \tag{10}$$

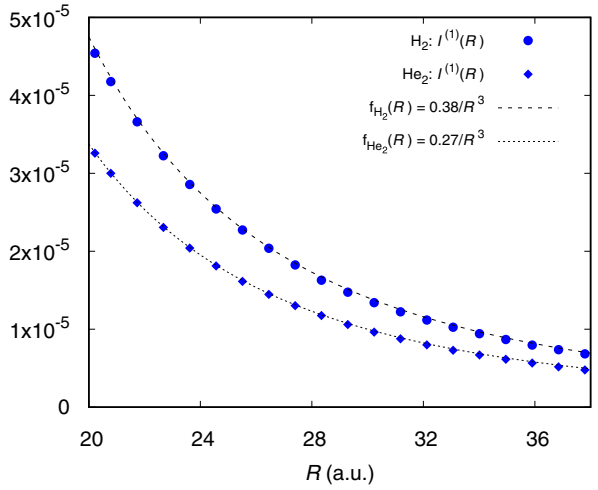


FIG. 2. $I^{(1)}(R)$ against R, the interatomic distance, for H_2 and He_2. The fitting corresponds to $f(R) = a/R^3$, where a is the fitted parameter and equals 0.38 and 0.27 for H_2 and He_2, respectively. In both cases the Pearson regression coefficient, r^2, is greater than 0.999. All quantities are in a.u.

To this aim, we have chosen two simple molecules, H_2 and He_2, and we have performed full-configuration interaction calculations[1] with the aug-cc-pVDZ basis set at different interatomic separations (R). We have computed Eq. (10) and plotted $I^{(1)}(R)$ against R (see Fig. 2).

The numerical results confirm the predicted $1/R^3$ dependency of the vdW correction to the intracule of the pair density. The fitting procedure employs the points presented in Fig. 2 and uses a least squares minimization analysis to determine the Pearson regression coefficient and the a parameter in the $f(R) = a/R^3$ fitting function.

The intracule pair density is the simplest quantity in terms of which we can express explicitly and exactly the electron-electron repulsion energy,

$$V_{ee} = \int du \frac{I(u)}{u}, \tag{11}$$

and the integration of the first-order correction to the intracule divided by the electron-electron distance provides the vdW energy:

$$V_{ee}^{(2)} = \int du \frac{I^{(1)}(u)}{u}. \tag{12}$$

[1]The calculations have been performed with a modified version of the program of Knowles and Handy [27] and the pair density matrices have been obtained with our in-house DMN code. [28] The intracule calculation employed the in-house RHO2_OPS [29] code, which uses the algorithm [30] of Cioslowski and Liu.

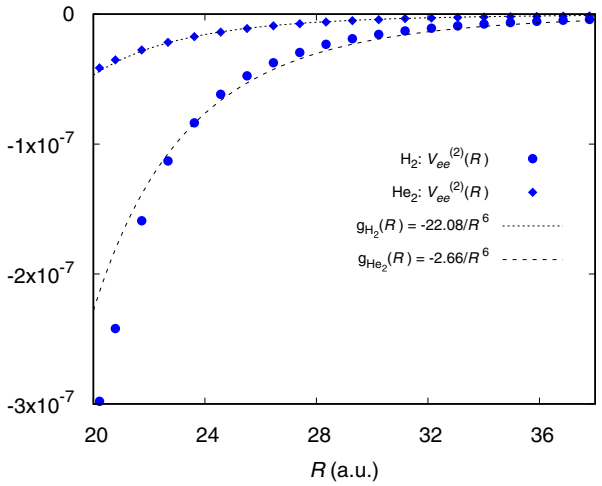


FIG. 3. $V_{ee}^{(2)}(R)$ against R, the interatomic distance, for H_2 and He_2. The fitting corresponds to $g(R) = a/R^6$, where a is the fitted parameter and equals -22.08 and -2.66 for H_2 and He_2, respectively. In both cases the Pearson coefficient, r^2, is greater than 0.97. All quantities are in a.u.

Through Eq. (12), we can calculate the vdW energy using the first-order correction to the intracule function. In Fig. 3 we observe that at large distances these functions reproduce the $1/R^6$ dependency. The fitting is not as good as in the intracule function because of the large power dependency, which increases the numerical error.

IV. CONCLUSIONS

We have unveiled a universal condition of the intracule of the pair density: the vdW contribution to the intracule of the pair density at R should decay as $1/R^3$, R being the separation of two fragments. This condition is connected to the well-known $1/R^6$ decay of the vdW energy and it can be recovered from the vdW contribution to the intracule of the pair density [see Eq. (12)]. This requirement is a salient signature of vdW interactions that can be employed as a stringent constraint in a judicious construction of new methods and approximations in electronic structure theory including vdW interactions [24–26]. Since the vdW correction to the intracule of the pair density shows a lower power dependency than the energy one, it is also less prone to numerical errors (compare fittings in Figs. 2 and 3).

ACKNOWLEDGMENTS

The authors thank Dr. Eloy Ramos-Cordoba for helpful discussions. This research has been funded by Spanish MINECO/FEDER Project No. CTQ2014-52525-P and the Basque Country Consolidated Group Project No. IT588-13. M.V.-N. and M.R.-M. acknowledge the Spanish Ministry of Economy, Industry and Competitiveness (MINECO) and the Spanish Ministry of Education, Culture and Sports for the doctoral Grants No. BES-2015-072734 and No. FPU-2013/00176, respectively.

[1] J. Hermann, R. A. DiStasio, Jr., and A. Tkatchenko, Chem. Rev. **117**, 4714 (2017).
[2] L. Yang, C. Adam, G. S. Nichol, and S. L. Cockroft, Nat. Chem. **5**, 1006 (2013).
[3] J. Černý, M. Kabelác, and P. Hobza, J. Am. Chem. Soc. **130**, 16055 (2008).
[4] J. Klimeš, D. R. Bowler, and A. Michaelides, Phys. Rev. B **83**, 195131 (2011).
[5] M. Rosa, S. Corni, and R. Di Felice, Phys. Rev. B **90**, 125448 (2014).
[6] F. W. DelRio, M. P. de Boer, J. A. Knapp, E. D. Reedy, P. J. Clews, and M. L. Dunn, Nat. Mater. **4**, 629 (2005).
[7] S. Kristyán and P. Pulay, Chem. Phys. Lett. **229**, 175 (1994).
[8] K. Berland, V. R. Cooper, K. Lee, E. Schröder, T. Thonhauser, P. Hyldgaard, and B. I. Lundqvist, Rep. Prog. Phys. **78**, 066501 (2015).
[9] T. Thonhauser, V. R. Cooper, S. Li, A. Puzder, P. Hyldgaard, and D. C. Langreth, Phys. Rev. B **76**, 125112 (2007).
[10] S. Grimme, WIREs, Comput. Mol. Sci. **1**, 211 (2011).
[11] A. D. Becke and E. R. Johnson, J. Chem. Phys. **123**, 154101 (2005).
[12] A. D. Becke and E. R. Johnson, J. Chem. Phys. **127**, 154108 (2007).
[13] K. Rapcewicz and N. W. Ashcroft, Phys. Rev. B **44**, 4032(R) (1991).
[14] P. Hyldgaard, K. Berland, and E. Schröder, Phys. Rev. B **90**, 075148 (2014).
[15] N. Ferri, R. A. DiStasio, Jr., A. Ambrosetti, R. Car, and A. Tkatchenko, Phys. Rev. Lett. **114**, 176802 (2015).
[16] N. Ferri, A. Ambrosetti, and A. Tkatchenko, Phys. Rev. Materials **1**, 026003 (2017).
[17] L. Pauling and E. B. Wilson, *Introduction to Quantum Mechanics* (Dover Publications, Inc., New York, 1935).
[18] A. Unsöld, Z. Phys. **43**, 563 (1927).
[19] J. O. Hirschfelder, W. B. Brown, and S. T. Epstein, Adv. Quantum Chem. **1**, 255 (1964).
[20] C. A. Coulson and A. H. Neilson, Proc. Phys. Soc. London **78**, 831 (1961).
[21] A. J. Thakkar, A. Tripathi, and V. H. Smith, Int. J. Quantum Chem. **26**, 157 (1984).
[22] A. J. Thakkar, A. N. Tripathi, and V. H. Smith, Jr., Phys. Rev. A **29**, 1108 (1984).
[23] M. Rodríguez-Mayorga, E. Ramos-Cordoba, M. Via-Nadal, M. Piris, and E. Matito, Phys. Chem. Chem. Phys. **19**, 24029 (2017).
[24] D. A. Mazziotti, Phys. Rev. Lett. **108**, 263002 (2012).
[25] A. W. Schlimgen, C. W. Heaps, and D. A. Mazziotti, J. Phys. Chem. Lett. **7**, 627 (2016).
[26] D. A. Mazziotti, Phys. Rev. Lett. **117**, 153001 (2016).
[27] P. J. Knowles and N. C. Handy, Comput. Phys. Commun. **54**, 75 (1989).
[28] E. Matito and F. Feixas, *DMn program*, 2009, University of Girona (Spain) and University of Szczecin (Poland).
[29] M. Rodríguez-Mayorga, *RHO2-OPS: 2-dm operations,* 2016, Institute of Computational Chemistry and Catalysis, University of Girona, Catalonia, Spain.
[30] J. Cioslowski and G. Liu, J. Chem. Phys. **105**, 4151 (1996).

Chapter 5

Benchmarking of reduced density matrix functional approximations in molecules

A density matrix functional approximations benchmark for atoms and molecules

Mireia Via-Nadal,[†,‡] Eloy Ramos-Cordoba,[†,‡] Pierre François Loos,[¶] and Eduard Matito[*,†,§]

E-mail: ematito@gmail.com

Abstract

At present, several functional approximations to reduced density matrix functional theory (rDMFT) have been developed, yet very few benchmarks exist in the literature. Usually, the energy is used as the main criterion for validating the performance of reduced density matrix functional approximations (rDMFAs). Besides the energy, more chemical properties should be considered, along with further analysis on the design of the expressions. In this paper, we report a systematic study of the performance of several rDMFAs on a benchmark set of first and second row molecules, which aims to complement an already published benchmark that used the harmonium atom ($N = 2$). Our benchmark indicates the need for considering other properties besides the energy in order to develop approximations.

Introduction

Because the electronic operators are one- and two-electron operators, the electronic energy can be expressed in terms of the first- and second-order reduced density matrices

(1-rDM and 2-rDM, respectively),

$$E = -\frac{1}{2}\int_{\mathbf{1}\to\mathbf{1}'} \nabla_{\mathbf{1}}^2 \rho_1(\mathbf{1}';\mathbf{1}) d\mathbf{1}' + \int \frac{\rho(\mathbf{1})}{|\boldsymbol{R}-\boldsymbol{r}_1|} d\mathbf{1} + \iint \frac{\rho_2(\mathbf{1},\mathbf{2})}{|\boldsymbol{r}_1-\boldsymbol{r}_2|} d\mathbf{1} d\mathbf{2}, \tag{1}$$

where we have expressed space and spin coordinates as boldface numbers, $\mathbf{1} \equiv \overrightarrow{r_1}, \sigma_1$. Because of this, an intuitive path for solving the Schrödinger equation implies using the 2-rDM of a system instead of an inaccessibly large wavefunction. This approach is treated in reduced density matrix functional theory (rDMFT),[1–3] whose grounds were introduced by Gilbert in 1975 as an extension to the Hohenberg-Kohn theorem for nonlocal external potentials.[1] Nowadays, several approximations to the exact functional have been developed in the literature, yet their application into actual quantum chemistry calculations is far from widespread. Albeit rDMFT approximations have been benchmarked already in the literature,[4–9] a broad and extensive study of the energy and other properties of a large collection of approximations in molecules is still missing.

The energy is an explicit functional of the 2-rDM by virtue of eq. 1. The 2-rDM is defined as

$$\rho_2(\mathbf{1}',\mathbf{2}';\mathbf{1},\mathbf{2}) = N(N-1)\int \Psi^*(\mathbf{1}',\mathbf{2}',\mathbf{3}\ldots\boldsymbol{N})\,\Psi(\mathbf{1},\mathbf{2},\mathbf{3}\ldots\boldsymbol{N})\,d\mathbf{3}\ldots d\boldsymbol{N}, \tag{2}$$

where we consider the McWeeny normalization factor to normalize the density matrix.[10] Besides its lower dimensionality with respect to the N-electron wavefunction, the advantage of working with density matrices is the statistical information of electron pair probabilities contained in the elements of the 2-rDM. The diagonal elements of the 2-rDM form the pair density, $\rho_2(\mathbf{1},\mathbf{2};\mathbf{1},\mathbf{2}) \equiv \rho_2(\mathbf{1},\mathbf{2})$. Integration of the 2-rDM by the convenient coordinates reduces to the 1-rDM,

$$\rho_1(\mathbf{1}';\mathbf{1}) = \frac{1}{N-1}\int \rho_2(\mathbf{1}',\mathbf{2}';\mathbf{1},\mathbf{2})\delta(\mathbf{2}-\mathbf{2}')d\mathbf{2}d\mathbf{2}', \tag{3}$$

which, in turn, bears the well-known electron density in its diagonal elements, $\rho_1(\mathbf{1};\mathbf{1}) \equiv \rho_1(\mathbf{1})$. Hence, the electronic energy of a system can be calculated utilizing the 2-rDM. Even though 2-rDMs are smaller objects than wavefunctions, their size may be impractical to store and manipulate. rDMFT proposes the reduction of the computational cost by means of building approximations to the 2-rDM, and construct them in terms of a known 1-rDM.

Because calculations are mostly performed using orbital basis sets, we consider the orbital representation of the 2-rDM, in which $\rho_2(\mathbf{1}',\mathbf{2}';\mathbf{1},\mathbf{2})$ is expanded in a set of orbitals $\{\phi_i(\mathbf{1})\}_{i=1,K}$,

$$\rho_2(\mathbf{1}',\mathbf{2}';\mathbf{1},\mathbf{2}) = \sum_{\substack{ij\\kl}}^{K} {}^2\mathrm{D}^{ij}_{kl}\phi^*_i(\mathbf{1}')\phi^*_j(\mathbf{2}')\phi_k(\mathbf{1})\phi_l(\mathbf{2}), \tag{4}$$

where ${}^2\mathrm{D}^{ij}_{kl}$ is a tensor of K^4 elements named 2-density matrix (2-DM). A similar expression can be written for the 1-density matrix (1-DM) where, when extended in the natural orbitals basis set, the 1-DM is diagonal and its elements are the natural occupancies,

$$\rho_1(\mathbf{1}';\mathbf{1}) = \sum_{ij}^{K} {}^1\mathrm{D}^i_j\phi^*_i(\mathbf{1}')\phi_j(\mathbf{1}) = \sum_{i}^{K} n_i\,\chi^*_i(\mathbf{1}')\chi_j(\mathbf{1}), \tag{5}$$

being $\{n_i\}_{i=1,K}$ the natural occupation numbers and $\{\chi_i\}_{i=1,K}$ the natural orbitals. In accordance with our previous study, we adopt the natural orbital representation of the 2-DM and, therefore, 2-DM approximations are constructed from natural occupation numbers. In this work, we construct the approximate 2-DMs utilizing the natural occupancies from a full configuration interaction (FCI) wavefunction.

A 2-DM must correspond to an N-particle fermionic wavefunction, according to eq. 2, in order to bear a variational electronic energy after optimization.[11] When a 2-DM can be mapped back to an antisymmetric N-wavefunction, such 2-DM is called N-representable. There exists a set of necessary requirements, known as N-representability conditions, that

a density matrix must fulfill to be N-representable. Even though the set of conditions that guarantee the N-representability of the 2-DM is known,[12] they are, in practice, impossible to impose because they require higher-order density matrices, which may not exist, or bear large computational cost. Instead, three necessary (yet not sufficient) probability-based conditions are usually imposed, since they only involve the manipulation of the 2-DM. These conditions are known as the P-, Q-, and G-conditions, and define the positive semidefinite character of P, Q and G matrices,[5,13–16]

$$P^{\sigma\sigma'}_{ij,kl} = \langle\Psi|a^{\dagger}_{i\sigma}a^{\dagger}_{j\sigma'}a_{k\sigma}a_{l\sigma'}|\Psi\rangle \tag{6}$$

$$Q^{\sigma\sigma'}_{ij,kl} = \langle\Psi|a_{i\sigma}a_{j\sigma'}a^{\dagger}_{k\sigma}a^{\dagger}_{l\sigma'}|\Psi\rangle \tag{7}$$

$$G^{\sigma\sigma'}_{ij,kl} = \langle\Psi|a_{i\sigma}a^{\dagger}_{j\sigma'}a^{\dagger}_{k\sigma}a_{l\sigma'}|\Psi\rangle \tag{8}$$

where $\hat{a}_{i\sigma}$ and $\hat{a}^{\dagger}_{i\sigma}$ are the annihilation and creation operators from the second quantization language.[17] Because the eigenvalues of these matrices must be positive semidefinite, such conditions are also known as positivity conditions.[18,19] The eigenvalues of the P, Q and G matrices refer to different probabilities of electron distributions among orbitals and, in the natural orbital representation, restrictions are imposed to occupancies in natural orbitals (particle and hole probabilities).[5] Indeed, the definition of the P matrix corresponds to the second quantization definition of the 2-DM and involves the semidefinite positive character of the geminal occupancies. Q and G conditions involve the hole $(1 - n_i)$ and particle-hole $n_i(1 - n_i)$ probabilities, respectively.

2-DM approximations

A group of 14 rDMFAs is evaluated in this study, which are the same ones considered in our former study.[8] Those functional approximations create elements that accompany the Coulomb (J), exchange (K) and time-inversion (L) two-electron integrals,[20] and, ac-

cording to the approach considered for constructing the approximation, rDMFAs can be classified in two groups: K-functionals and JKL-functionals. The first group adapts the Hartree–Fock (HF) expression for the Coulombic terms and propose an expression to model the exchange elements of the 2-DM:[5,21,22]

$$^2\mathrm{D}^{\mathrm{X},\alpha\beta}_{ij,kl} = n^\alpha_i n^\beta_j \delta_{ik}\delta_{jl}, \tag{9}$$

$$^2\mathrm{D}^{\mathrm{X},\alpha\alpha}_{ij,kl} = n^\alpha_i n^\alpha_j \delta_{ik}\delta_{jl} - f^X(n^\alpha_i, n^\alpha_j)\delta_{il}\delta_{jk}, \tag{10}$$

where $f^{\mathrm{X}}(n^\alpha_i, n^\alpha_j)$ denotes the exchange functional. Because the approximations only differentiate among them by the exchange elements, we refer to them as K-functionals. The expressions taken by each 2-rDM approximation are summarized in Table 1. The most elemental approximation to the 2-DM is taking the actual exchange expression from HF,[23] which leads to the SD approximation. The SD 2-DM was originally not designed as an actual functional approximation within rDMFT because, upon optimization of natural orbitals and occupancies, it reduces to the actual HF results. Because the HF expression is used but neither the HF occupancies nor orbitals, the SD 2-DM represents an approximation to the 2-DM by itself.

The first actual functional approximation was MBB, derived independently by Müller by one side, and by Buijse and Baerends by the other one.[21,22,24] MBB was derived from the requirement of minimal deviation of the Pauli principle condition. On the other hand, Goedecker and Umrigar designed the GU approximation to correct the self-interacting error in MBB.[25] Gritsenko, Pernal, and Baerends suggested three versions to improve the MBB functional: BBC1, BBC2, and BBC3.[26] In this work, we consider BBC2, which uses the Fermi level F_L to introduce electron pair interactions according to the orbitals in which they are found. BBC3 requires the identification of the type of orbital and thus we cannot construct the 2-DM only with the FCI occupancies.[26]

Table 1: Exchange expressions $f^{\mathrm{X}}(n_i, n_j)$ (eq. 10) per each K-functional studied in this work.

Functional	$f^{\mathrm{X}}(n_i, n_j)$		
SD	$n_i n_j$	$\forall i, j$	
MBB	$(n_i n_j)^{1/2}$	$\forall i, j$	
GU	$(n_i n_j)^{1/2}$	$i \neq j$	
	$n_i n_j$	$i = j$	
BBC2	n_i	$i = j$	
	$n_i n_j$	$i \neq j \wedge i \in [1; F_L] \wedge j \in [1; F_L]$	
	$-(n_i n_j)^{1/2}$	$i \neq j \wedge i \in (F_L; \infty) \wedge j \in (F_L; \infty)$	
	$(n_i n_j)^{1/2}$	otherwise	
CA	$[(n_i(1-n_i)n_j(1-n_j)]^{1/2} + n_i n_j$	$\forall i, j$	
CGA	$\frac{1}{2}\left[\left[(n_i(2-n_i)n_j(2-n_j)\right]^{1/2} + n_i n_j\right]$	$\forall i, j$	
ML	$n_i n_j(a + b\, n_i n_j)(1 + c\, n_i n_j)$	$\forall i, j$	$a = 126.3101$ $b = 2213.33$ $c = 2338.64$
MLSIC	$n_i n_j(a + b\, n_i n_j)/(1 + c\, n_i n_j)$	$i \neq j$	$a = 1298.78$
	$n_i n_j$	$i = j$	$b = 35114.4$ $c = 36412.2$

The CA and CGA functionals were proposed by Csányi and Arias,[27,28] and are based on a tensor product expansion of the 2-rDM in order to reproduce the homogeneous electron gas. CGA was later presented as an improvement of CA.[28] The functionals derived

by Marques and Lathiotakis, the ML and MLSIC functional approximations, replace the square root in MBB with a Padé approximant, and include three parameters. MLSIC is the self-interaction-corrected version of ML.[29] Notice that both ML and MLSIC functionals were developed exclusively to reproduce the total energy. It is our prerogative to use them in this study to calculate other properties assuming their expressions are general enough for this purpose.

On the other hand, the second group of rDMFAs studied in this work are the *JKL*-functionals, based on the spectral expansion of the cumulant of the 2-DM.[30] The cumulant, usually labelled as ${}^2\Gamma^{kl}_{ij}$ in the orbital representation, contains the missing correlation effects that the SD approximation of the 2-DM does not include. In other words, it contains all the electron correlation beyond HF in the 2-DM,

$$ {}^2\mathrm{D}_{ij,kl} = {}^2\mathrm{D}^{\mathrm{SD}}_{ij,kl} + {}^2\Gamma_{ij,kl}. \tag{11} $$

Because the actual expression of the cumulant in terms of natural occupancies is not known, the PNOFs (after "Piris Natural Orbital Functionals") aim to approximate the 2-DM via reconstructing the cumulant. To reduce the four-index dimensionality, PNOFs assume that the cumulant matrix is quite sparse and approximate it by splitting it into two auxiliary matrices, known as Δ and Π,[30]

$$ {}^2\Gamma^{\alpha\beta}_{ij,kl} = -\Delta^{\alpha\beta}_{ij}\delta_{ik}\delta_{jl} + \Pi_{ik}\delta_{ij}\delta_{kl}, \tag{12} $$

$$ {}^2\Gamma^{\alpha\alpha}_{ij,kl} = -\Delta^{\alpha\alpha}_{ij}\left(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}\right). \tag{13} $$

In contrast to the *K*-functionals, the Coulomb elements are also approximated. According to the expressions granted to Δ and Π, different versions of PNOFs exist, named PNOF*i* with $i = 2\ldots,7$.[31–38] Table 2 gathers the expressions for Δ and Π matrices for all PNOFs.

Table 2: Expressions for the Δ and Π matrix elements for the PNOFi, $i = 1 \ldots 7$, according to eq. 13. The diagonal elements for any PNOFi are $\Delta_{ii} = n_i^2$ and $\Pi_{ii} = n_i$. Appearing elements in the Table are $h_i = 1 - n_i$, $T_{ij} = n_i n_j - \Delta_{ij}$, and $S_F = \sum_{i=1}^{F_L} h_i$. For PNOF6$x$, x = d, u, h, where each ingredient used is defined through eqs. 14–18. For PNOF5 and PNOF7, Ω_g is the subspace containing orbital g (which is imposed to be below F_L) and several orbitals above F_L.

Functional	Δ_{ij}	Π_{ij}	
PNOF2	$h_i h_j$	$\sqrt{n_i n_j} + \sqrt{h_i h_j} + T_{ij}$	$i \wedge j \in [1, F_L]$
	$h_i n_j \left(\frac{1-s_F}{s_F}\right)$	$\sqrt{n_i n_j} - \sqrt{h_i n_j} + T_{ij}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i h_j \left(\frac{1-s_F}{s_F}\right)$	$\sqrt{n_i n_j} - \sqrt{n_i h_j} + T_{ij}$	$i \in (F_L, K] \wedge j \in [1, F_L]$
	$n_i n_j$	T_{ij}	$i \wedge j \in (F_L, K]$
PNOF3	$h_i h_j$	$n_i n_j - \sqrt{n_i n_j}$	$i \wedge j \in [1, F_L]$
	$h_i n_j \left(\frac{1-s_F}{s_F}\right)$	$n_i n_j - \sqrt{n_i n_j} - \sqrt{h_i n_j}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i h_j \left(\frac{1-s_F}{s_F}\right)$	$n_i n_j - \sqrt{n_i n_j} - \sqrt{n_i h_j}$	$j \in [1, F_L] \wedge i \in (F_L, K]$
	$n_i n_j$	$n_i n_j + \sqrt{n_i n_j}$	$i \wedge j \in (F_L, K]$
PNOF4	$h_i h_j)$	$-\sqrt{h_i h_j}$	$i \wedge j \in [1, F_L]$
	$h_i n_j \left(\frac{1-s_F}{s_F}\right)$	$-\sqrt{\left(\frac{h_i n_j}{S_F}\right)\left(n_i - n_j + \frac{h_i n_j}{S_F}\right)}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i h_j \left(\frac{1-s_F}{s_F}\right)$	$-\sqrt{\left(\frac{n_i h_j}{S_F}\right)\left(-n_i + n_j + \frac{n_i h_j}{S_F}\right)}$	$i \in [1, F_L] \wedge j \in (F_L, K]$
	$n_i n_j$	$\sqrt{n_i n_j}$	$i \wedge j \in (F_L, K]$
PNOF5	$n_i n_j$	$-\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i = g \vee j = g)$
	$n_i n_j$	$\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i \wedge j \in (F_L, K])$
PNOF6x	$h_i h_j e^{-2S_F}$	$-e^{-S_F}\sqrt{h_i h_j}$	$i \wedge j \in [1, F_L]$
	$\gamma_i \gamma_j / S_\gamma^x$	$-\sqrt{\left[n_i h_j + \frac{\gamma_i \gamma_j}{S_\gamma^x}\right]\left[h_i n_j + \frac{\gamma_i \gamma_j}{S_\gamma^x}\right]}$	$i \in [1, F_L] \wedge j \in (F_L, K]$ $\vee\ j \in [1, F_L] \wedge i \in (F_L, K]$
	$n_i n_j e^{-2S_F}$	$e^{-S_F}\sqrt{n_i n_j}$	$i \wedge j \in (F_L, K]$

PNOF7	$n_i n_j$	$-\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i = g \vee j = g)$
	$n_i n_j$	$\sqrt{n_i n_j}$	$(i \wedge j \in \Omega_g) \wedge (i \wedge j \in (F_L, K])$
	0	$-\sqrt{n_i h_i n_j h_j}$	$(i \vee j) \in [1, F_L]$ $\wedge \left[(i \in \Omega_g \wedge j \notin \Omega_g) \vee (j \in \Omega_g \wedge i \notin \Omega_g)\right]$
	0	$\sqrt{n_i h_i n_j h_j}$	$(i \wedge j) \in (F_L, \infty)$ $\wedge \left[(i \in \Omega_g \wedge j \notin \Omega_g) \vee (j \in \Omega_g \wedge i \notin \Omega_g)\right]$

The shape of S_γ^x in PNOF6 depends on the orbitals considered to construct it, which can be either the orbitals placed above (x = u, up) or below (x = d, down) the F_L,

$$S_\gamma^{\mathrm{d}} = \sum_{i=1}^{F_L} \gamma_i; \tag{14}$$

$$S_\gamma^{\mathrm{u}} = \sum_{i>F_L}^{K} \gamma_i. \tag{15}$$

An average of both cases results in PNOF6h (x = h, half),

$$S_\gamma^{\mathrm{h}} = (S^{\mathrm{d}} + S^{\mathrm{u}})/2. \tag{16}$$

The γ_i term takes the following expression,

$$\gamma = n_i h_i + \kappa_i^2 - \kappa_i \sum_{i=j}^{F_L} \kappa_j, \tag{17}$$

where

$$\kappa_i = \begin{cases} h_i e^{-S_F} & \text{when } i \in [1, F_L] \\ n_i e^{-S_F} & \text{when } i \in (F_L, K]. \end{cases} \tag{18}$$

A perfect pairing approach was initially imposed in PNOF5, PNOF6, and PNOF7, in which the occupancies of two orbitals i and j are forced to be coupled, $n_i + n_j = 1$, $i \neq j$, being i an orbital below the F_L and j above it, or viceversa.[34,35] This perfect pairing ap-

proach is an unphysical condition to the construction of the rDMFT approximation. In this benchmark, as we are using the FCI occupancies, this restriction cannot be fulfilled. Instead, we consider the extended versions of these functionals, in which the occupancy of an orbital below the F_L is coupled to the occupancies of a set of orbitals above it and, hence, both PNOFs are free of restrictions in a closed-shell system.[37,38] In this study, we have not considered PNOF5 and PNOF7 due to the imposibility to define the Ω_g subspaces, which are automatically chosen in the optimization algorithm of the approximations.

Because the 2-DM approximations are built using the exact (within the basis set) natural orbitals and occupation numbers and not optimized for each approximation, the errors analyzed from each test only consider the effect produced by the expression designed to approximate the 2-DM. In a sense, the accuracy of the expressions for the 2-DM approximations is evaluated. This type of error is sometimes called functional–driven error in KS DFT,[39] which in that framework evaluates the true error in an exchange-correlation approximation of any given density. Let us introduce the total energy error in rDMFT:

$$\mathrm{E_{total}} = E\left[\{n_i,\chi_i\}\right] - E^X\left[\left\{n_i^X,\chi_i^X\right\}\right], \tag{19}$$

where the first term is the exact energy and the second term is the optimized rDMFA energy. This error can be split in the functional–driven and the 1-DM–driven error, $\mathrm{E_{total}} = \mathrm{E_F} + \mathrm{E_{DM}}$. In the rDMFT framework, the energy functional–driven error is:

$$\mathrm{E_F} = E\left[\{n_i,\chi_i\}\right] - E^X\left[\{n_i,\chi_i\}\right], \tag{20}$$

in which the second term in the *r.h.s.* of eq. 20 is the energy of the rDMFA obtained with the FCI natural occupancies and orbitals – that is, a non-optimized energy. On the other

hand, the 1-DM–driven error (the analogy to the density–driven error in KS DFT)[39] is:

$$\mathrm{E_{DM}} = E^X\left[\{n_i, \chi_i\}\right] - E^X\left[\left\{n_i^X, \chi_i^X\right\}\right], \tag{21}$$

which is also the self-consistent error of the approximation. In this work, the 1-DM–driven error is not analyzed.

Computational details

The physical basis and quality of the 2-DM approximations are evaluated through a set of diatomic molecules composed of first- and second-row elements. The molecules considered and their interatomic distances are collected in Table 3. This test considers the experimental equilibrium geometry R_{eq}, and a stretched geometry at $R_{\mathrm{diss}} = 5R_{\mathrm{eq}}$. Besides this set of molecules, this work also analyzes the performance in predicting the Be_2 energy curve,[40–42] the size consistency of the approximations via $N = 4, 6, 8$-vertex polyhedra resulting from N hydrogen atoms separated 10 Å from the geometric center of the respective polyhedron,[43,44] and different correlation regimes through the D_{4h}/D_{2h} potential energy surface of H_4.[44–47] The geometrical parameters for these four systems are summarized in Table 4.

Table 3: Experimental equilibrium geometries for the diatomic molecules in Angstroms, taken from the NIST database.[48] [a]Equilibrium distance reference is CCSD(T)/aug-cc-pVQZ from the NIST database.[48] [b]Equilibrium distance reference is CCSD(T)/aug-cc-pVQZ from the NIST database.[48]

Molecule	R_{eq}	Molecule	R_{eq}
H_2	0.741	LiH	1.595
He_2	2.963[a]	^{2}BeH	1.343
Li_2	2.673	^{2}CH	1.120
Be_2	2.460	NH	1.034
C_2	1.243	^{2}OH	0.970
N_2	1.098	HF	0.917
$^3O_2$	1.208	LiF	1.564
$^1O_2$	1.218[b]	CO	1.128
F_2	1.412	^{2}NO	1.154
		^{2}CN	1.172

Table 4: Geometries for the non-diatomic molecules considered in this test, in Angstroms and degrees. R defines the distance between the geometric center and a hydrogen placed in a vertex.[44] In H_4, θ indicates the angle formed between the geometric center and two adjacent H atoms, $\angle$H0H.[46]

Molecule	Geometries
H_4	$R = 0.8$, $\theta = 70$ (D_{2h})
	$R = 4.0$, $\theta = 70$ (D_{2h})
	$R = 0.8$, $\theta = 90$ (D_{4h})
	$R = 4.0$, $\theta = 90$ (D_{4h})
H_n polyhedra	$N = 2$ (hydrogen molecule)
	$N = 4$ (tetrahedron)
	$N = 6$ (octahedron)
	$N = 8$ (cube)
	$R = 10.0$

The reference 2-DM and natural occupancies were obtained with the in-house DMN code,[49] which uses the CI vector obtained from a selected configuration interaction (sCI) based on the CIPCI algorithm[50–52] to construct the reference density matrices. The matrix elements of the approximated 2-DMs, $^2D^X$, are constructed on the fly per each test used, and are generated with the reference natural orbitals and occupation numbers, $n_i\chi_i \rightarrow {}^2D^X$, which are considered to be exact within the given basis set. DIs are computed with the in-house codes RHO−OPS[53] and APOST−3D.[54] Intracule radial probability distributions

are obtained with the in-house code RHO2−OPS[55] that uses the algorithm proposed by Cioslowski and Liu .[56]

Tests

In the present work, we analyze the physical background of the 2-DM approximations and their performance in predicting energies and chemical properties. We evaluate the rDMFAs with:

a) The error in the trace of the approximated 2-DM,

$$\mathrm{E_{trace}} = \mathrm{Tr}\left[{}^2\mathrm{D}^{\mathrm{X}}\right] - \mathrm{Tr}\left[{}^2\mathrm{D}\right], \tag{22}$$

where the trace of the exact 2-DM,

$$\mathrm{Tr}\left[{}^2\mathrm{D}\right] = \sum_{ij}^{K} {}^2\mathrm{D}_{ij,ij}, \tag{23}$$

sums up to the number of electron pairs $N(N-1)$, according to the McWeeny normalization factor described in eq. 2.

b) The cumulative absolute error (CAE) in both the diagonal and all the elements of the approximated 2-DM,

$$\mathrm{CAE_{diag}} = \sum_{ij}^{K} |{}^2\mathrm{D}^{\mathrm{X}}_{ij,ij} - {}^2\mathrm{D}_{ij,ij}|; \tag{24}$$

$$\mathrm{CAE} = \sum_{ijkl}^{K} |{}^2\mathrm{D}^{\mathrm{X}}_{ij,kl} - {}^2\mathrm{D}_{ij,kl}|. \tag{25}$$

Both CAE are normalized by the number of electron pairs for each molecule

$$\mathrm{CAE'_{diag}} = 2\mathrm{CAE_{diag}}/N(N-1) \tag{26}$$

$$\mathrm{CAE'} = 2\mathrm{CAE}/N(N-1) \tag{27}$$

and the average for all the molecules analyzed is presented in the Results section.

c) The evaluation of the eigenvalues of the positivity conditions (eqs. 6–8) by summing all the negative eigenvalues produced by the respective matrices.[8] For each molecule, this sum is normalized by the number of electron pairs to make all cases comparable.

d) The antisymmetry of the elements of the 2-DM,

$$\mathrm{E_{antisym}} = \sum_{ijkl,\sigma}^{K} |{}^2\mathrm{D}^{\mathrm{X},\sigma\sigma}_{ij,kl} + {}^2\mathrm{D}^{\mathrm{X},\sigma\sigma}_{ij,lk} + {}^2\mathrm{D}^{\mathrm{X},\sigma\sigma}_{ji,lk} + {}^2\mathrm{D}^{\mathrm{X},\sigma\sigma}_{ji,kl}|, \tag{28}$$

since the 2-DM must describe the interactions arising from the fermionic nature of electrons. The sum is normalized to the number of electron pairs of the 2-DM of each molecule.

e) The difference between the exact and the approximated delocalization index (DI)

$$\mathrm{E_{DI}} = \delta^{\mathrm{X}}(A,B) - \delta(A,B), \tag{29}$$

where the DI is defined as

$$\delta(A,B) = -2\int_A\int_B d\mathbf{1}d\mathbf{2}\left[\rho_2(\mathbf{1},\mathbf{2}) - \rho(\mathbf{1})\rho(\mathbf{2})\right]. \tag{30}$$

$\delta(A,B)$ is a measure of the bond order between two fragments within a molecule by examining the electron population between regions A and B.[57–59] Note that the elec-

tron density in eq. 30 is the same as the reference, FCI density, as in this work we use the exact natural orbitals and occupancies to construct the approximate 2-DM. Then, the test actually measures the accuracy of the approximation in predicting the number of electron pairs in both regions;

f) The error in the (radial) intracule probability density (IPD)

$$\mathrm{E}_{I(s)} = I(\rho_2, s) - I(\rho_2^{\mathrm{X}}, s), \tag{31}$$

where the IPD is an electron-pair distribution function that indicates the probability of two electrons being separated a given interelectronic distance s,[10,60,61]

$$I(s) \equiv I(\rho_2, s) = s^2 \int d\Omega_s d\mathbf{1} d\mathbf{2}\, \rho_2(\mathbf{1}, \mathbf{2}) \delta(\boldsymbol{s} - \boldsymbol{r_{12}}), \tag{32}$$

with $d\Omega_s = \sin\theta d\theta d\phi$ being the solid angle over the spherical coordinates, providing an spherical average of the probability density over space, and $\boldsymbol{r}_{12} = \boldsymbol{r}_2 - \boldsymbol{r}_1$.[61–63] Then, the error measured in this test, $\mathrm{E}_{I(s)}$, is a function that evaluates the ability of each functional approximation to reproduce all the $I(s)$ points. Notice that the IPD can be built from any 2-DM (for instance, the HF 2-DM, or an approximate 2-DM). In the specific case in which X = HF, the definition of Coulson's Coulomb hole is recovered in eq. 31.[64]

g) The relative error in the average distance between electron pairs $\langle s \rangle$ and its variance $\sigma^2 = \langle s^2 \rangle - \langle s \rangle^2$,

$$\mathrm{E}_{\langle s \rangle} = (\langle s \rangle_{\mathrm{X}} - \langle s \rangle)/\langle s \rangle \tag{33}$$

and

$$\mathrm{E}_{\sigma^2} = (\sigma^2_{\mathrm{X}} - \sigma^2)/\sigma^2, \tag{34}$$

$\langle s \rangle$ and σ^2 being properties that can be extracted from the IPD (eq. 32),

$$\langle s \rangle = \int ds\, I(s) s. \tag{35}$$

h) The relative error in interelectronic repulsion energy V_{ee},

$$\mathrm{E}_{V_{ee}} = \left(V_{ee}^{\mathrm{X}} - V_{ee} \right) / V_{ee}, \tag{36}$$

which can be calculated from the IPD,

$$V_{ee} = \int ds \frac{I(s)}{s}. \tag{37}$$

Recall that, in rDMFT, the only term that needs to be approximated from the electronic Hamiltonian is the electron-electron repulsion energy V_{ee}, so evaluating V_{ee} is equivalent to evaluate the functional–driven total energy error (Eq. 20).

i) The error in dissociation energies of the diatomic molecules,

$$\mathrm{E}_{D_e} = D_e^{\mathrm{X}} - D_e, \tag{38}$$

where we only consider V_{ee},

$$D_e := V_{ee}\left(50 R_{\mathrm{eq}}\right) - V_{ee}\left(R_{\mathrm{eq}}\right). \tag{39}$$

j) The size-extensivity of the approximations using the polyhedron-shaped H_N molecules. The increase in the number of hydrogens N permits the study of the linear behaviour of the energy with N. The difference between the approximate and exact V_{ee}'s is compared for each H_N. The error analyzed is

$$\mathrm{E}_{V_{ee}/N} = \left(V_{ee}^{\mathrm{X}} - V_{ee} \right) / N. \tag{40}$$

k) The behavior of the approximations under the effect of different regimes of electron correlation. H_4 in D_{2h}/D_{4h} point symmetry group permits the modulation of the type and amount of correlation present in the system via small geometrical modifications. Nondynamic correlation dominates the correlated motion of electrons when atoms are largely separated, or when the orbitals are degenerate (when $\theta = 90°$ and $R = 0.8$ Å). Instead, dynamic correlation overtakes nondynamic correlation when the H atoms are close to each other and there is no degeneracy in the molecular orbitals.[45–47] The error is calculated with respect to the interelectronic repulsion energy of the ground state geometry $V_{ee,\,\mathrm{ref}}$ (R = 0.8 Å, $\theta = 70°$),

$$\mathrm{E}_{\mathrm{H_4,corr}} = (V_{ee}^{\mathrm{X}} - V_{ee,\,\mathrm{ref}}^{\mathrm{X}}) - (V_{ee}^{\mathrm{FCI}} - V_{ee,\,\mathrm{ref}}^{\mathrm{FCI}}) \tag{41}$$

l) The long-range behavior of the approximated 2-DM to describe dispersion interactions. Our group recently discovered a footprint of London dispersion forces in the LR region of the IPD that behaves very similarly to the very well-known energetic decay of the energy.[44,65] The first-order correction of the IPD with the interfragment distance R behaves as $I^{(1)}(\rho_2, R) \propto R^{-}3$. This behavior can be obtained by subtracting the IPD built with the SD 2-DM from the exact one (that is, $I(\rho_2, R) - I(\rho_2^{\mathrm{SD}}, R)$) because this difference captures the dynamic correlation of the pair density.[44] Hence, their ability to describe dispersion interactions will be analyzed with

$$\mathrm{E}_{\mathrm{disp}}(R) = \left(I(\rho_2^{\mathrm{X}}, R) - I(\rho_2^{\mathrm{SD}}, R)\right) R^3 \tag{42}$$

Because the tests have mostly been applied to a set of diatomic molecules to evaluate the accuracy of the functionals, the errors produced by a rDMFA in a benchmarking test are averaged using root mean squared errors (RMSE) and mean signed deviations (MSD),

$$\mathrm{RMSE}_{\mathrm{rDMFA}} = \frac{\sum_i^{\mathrm{T}} \sqrt{\mathrm{E}_{i,\mathrm{rDMFA}}^2}}{N_{\mathrm{T}}}; \tag{43}$$

$$\mathrm{MSD}_{\mathrm{rDMFA}} = \frac{\sum_i^{\mathrm{T}} \mathrm{E}_{i,\mathrm{rDMFA}}}{N_{\mathrm{T}}}, \tag{44}$$

where N_{T} is the number of diatomic molecules in the set, and E_i is the error obtained in a certain test for a given rDMFA (eqs. 22, 29, 33, 34, 36, and 38).

Physical fundaments analysis of the approximate 2-DMs

Trace

By definition, the trace of the 2-DM equals the number of electron pairs in the system (see eq. 2). MBB, BBC2, CA, CGA, PNOF2, and PNOF4 were designed to reproduce the correct trace and thus their error is zero. Figure 1 contains the root mean squared error (RMSE) of the error in the traces computed for all the molecules considered. The SD, GU, MLSIC, and PNOF3 approximations produce the same traces due to the definition of the diagonal elements used to construct the approximation (n_i^2). This group appears under the SD label in Figure 1. This group of rDMFAs produces the largest errors overall. In contrast, ML gives incorrect traces due to its parametrization, with errors only slightly smaller than the SD group. PNOF6d and PNOF6u generate very accurate traces with small deviations, and their combination to generate the PNOF6h functional results into the exact trace.

The error committed in $5R_{\mathrm{eq}}$ geometries is larger than in R_{eq}, excepting the beryllium and helium molecules. The error trend in both cases correlates with the idempotency deviation of the 1-DM used to build the approximation, because for many functionals (the SD group) it coincides with the error on the trace except for a constant, $\sum_i n_i^2 - N(N-1)$ *vs* $\sum_i n_i(1-n_i) = N - \sum_i n_i^2$. The error in He_2 is the same for both geometries, whereas in Be_2 smaller errors are obtained for the stretched geometry. Similar trends are observed according to the idempotency deviations. In general, deviations in open-shell molecules are smaller than in the closed-shell cases. However, the magnitude of the errors for both cases

are coincident in $5R_{eq}$ geometries. The traces produced by K-functionals in closed-shell systems are larger than the expected ones, whereas some PNOF deviations are negative (mainly in PNOF6d but also in some PNOF6u traces obtianed). The trends obtained for this set of atoms are in concordance with the ones obtained in the harmonium atom.[8]

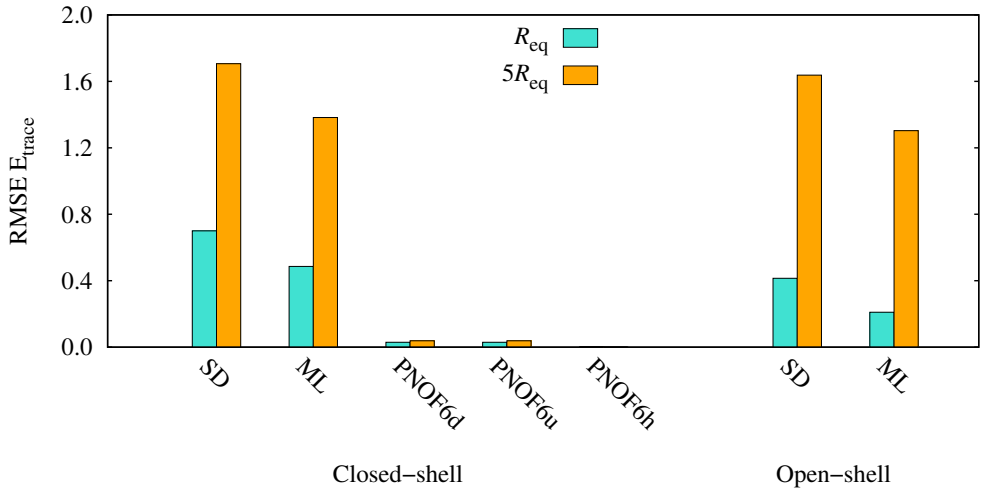


Figure 1: Root mean squared error (RMSE) of the trace error E_{trace} (eq. 22), produced by the SD group (including SD, GU, MLSIC, and PNOF3), the ML approximation, and the three PNOF6 variants considered in this work.

Elements of the 2-DM

Diagonal elements

In this section, we compare the approximate diagonal elements with the ones from the FCI 2-DM. The CAE for the diagonal elements is calculated by summing the absolute differences between the approximate and the FCI 2-DM (see eq. 24), and normalized by the number of electron pairs to make the error comparable (eq. 26).

MBB, BBC2, CA, and CGA present no error in the trace since they were designed to provide the exact one. However, their CAE, which is the same for the four approximations, is the largest among the K-functionals. These approximations provide the correct trace thanks to some unphysical elements in the 2-DM responsible for the self-interactions

($D^{\sigma\sigma}_{ii,ii} \neq 0$). SD, MLSIC, and GU are built using the same diagonal elements and, therefore, also produce the same CAE, which is the smallest among the K-only rDMFAs. It is interesting to remark that this latter group of rDMFAs presents a very simple form for the diagonal elements, $f(n_i, n_i) = n_i n_i$, which bears the smallest deviation within the K-functionals. On the other hand, the PNOF family of approximations generate diagonal elements that approximate the reference 2-DM elements, producing CAEs smaller than the K-functionals. PNOF2 and PNOF4 produce the correct trace, yet their elements (which ara coincident) bear some CAE. However, the magnitude of the CAE is minimal, and cannot be compared with the deviation produced by the group involving MBB, BBC2, CA, and CGA, which also produces the correct trace. PNOF3 presents the largest deviations among the PNOFs. The three versions of PNOF6 generate the smallest deviations in general. This trend is in agreement with the results obtained for the harmonium atom.[8]

The CAE of the diagonal elements of the K-functionals for molecules at $5R_{\text{eq}}$ is six times greater than the one produced at R_{eq}, and three times larger than in the PNOF family. The increase of electron correlation caused by the bond stretching is clearly problematic for the elements in the matrices. The diagonal elements in open-shell molecules are better approximated than in the closed-shell set of molecules. In this case, K-functionals produce CAEs of the same magnitude as the PNOFs.

It is worthy to mention the specific case of the hydrogen molecule, where the PNOF family (with exception of PNOF3) produces CAEs really close to zero. More precisely, the diagonal elements of PNOF2 and PNOF4 are equal to the reference ones in the stretched geometry of H_2 (see the Supporting Information). Indeed, PNOF2 and PNOF4 reduce to the exact functional for a two-electron system.[66] Same results to the CAE test were obtained for the harmonium atom.[8]

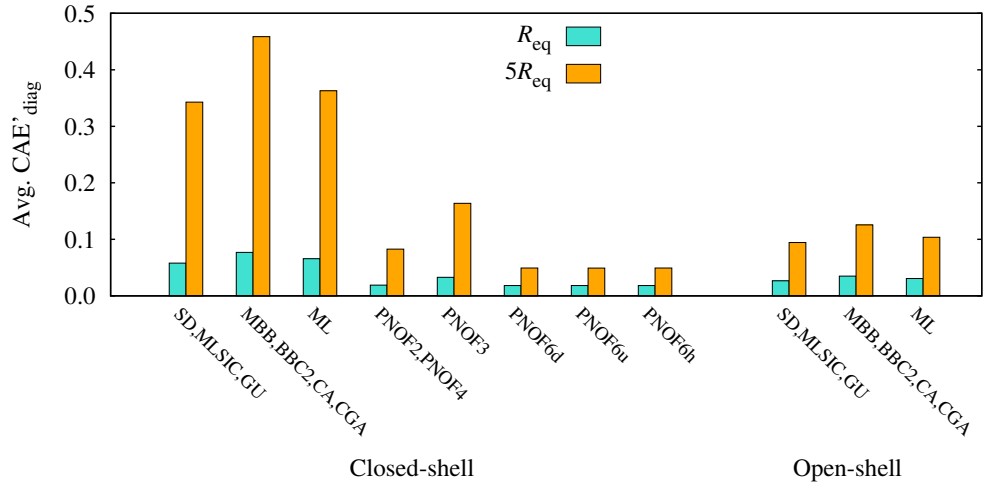


Figure 2: Average of the cumulative absolute errors (CAE), normalized to the number of electron pairs of each molecule $N(N-1)/2$, for the diagonal approximate elements of the 2-DM (Eq. 26).

Total CAE

Given a 1-rDM, all the components of the electronic Hamiltonian but V_{ee} are known (eq. 1) and, to approximate V_{ee}, one needs a 2-PD instead of the complete 2-rDM. The 2-rDM is, however, needed in order to compute other properties besides the energy. The average of the CAE for all the elements in the 2-DM (including the diagonal elements) are gathered in Figure 3.

According to Tables 1 and 2, each approximation has its expressions to construct the complete 2-DM, and hence none of them has the same elements as it occurred in the diagonal case. Because there are more elements to evaluate in this test, the dimension of the CAEs is larger than the CAE_{diag}. In contrast, the magnitude difference between equilibrium and stretched geometries is shortened with respect to the CAE_{diag} test: generally, greater CAE are obtained at $5R_{\text{eq}}$, yet the amount of deviations at equilibrium geometries is approximately equal. MBB, BBC2, CA, and CGA are the ones that produce larger differences between both geometries. SD, CA, and ML are the K-functionals with the lowest CAE, being SD the approximation that presents the most similar 2-DM elements with the reference 2-DM. PNOF2 and PNOF3 CAEs are a bit smaller than the SD and

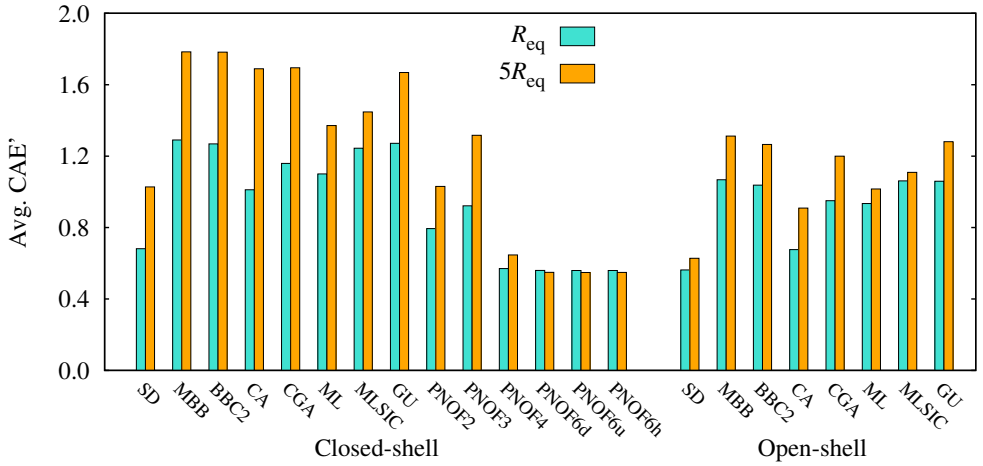


Figure 3: Average of the cumulative absolute errors (CAE), normalized to the number of electron pairs of each molecule $N(N-1)/2$, for all the elements of the approximate 2-DM (Eq. 27).

CA CAEs. PNOF4 and the three PNOF6 versions present the most consistent CAEs for both geometries and the lowest deviations overall for closed-shell molecules. Concerning the open-shell set, rDMFAs show similar CAEs for both geometries, but the equilibrium geometry is better approximated. Surprisingly, SD is the best rDMFA for this set.

The trends obtained in this section are in good agreement with the ones obtained for the diagonal elements alone, indicating that the main source of error in the approximate 2-DM elements comes from the diagonal components. However, the CA approximation presents the smallest increase of CAE compared to the $\mathrm{CAE}_{\mathrm{diag}}$, and very close to the SD CAE, a fact that was already observed in the harmonium atom in non-strongly correlated regimes.[8] The elements of the CA 2-DM are better approximated than its updated version, CGA. Despite the off-diagonal elements of ML and MLSIC were not specially designed to reproduce the exact elements, their CAEs are not the largest obtained in this test, and MBB, BBC2, CGA, and GU are the ones with the least accurate elements in the 2-DM. As in the $\mathrm{CAE}_{\mathrm{diag}}$ test, the CAE obtained for the PNOFs (excepting PNOF3) in H_2 is really close to zero, where PNOF4 and PNOF6 produce the closest 2-DM to the exact one.

Positivity conditions

The semidefinite positive character of the eigenvalues for the P, Q and G matrices are evaluated to determine wether the rDMFAs fulfill some of the known N-representability conditions of the 2-DM. We have analyzed the sum of all negative eigenvalues of those matrices, and divided them by the number of elements in the 2-DM to make all the systems comparable.

The PNOF6d approximation is the only functional that fulfills the three conditions for any system considered, whereas PNOF6u and, consequently, PNOF6h generate a small sum of negative eigenvalues in the P and Q matrices. In the harmonium case presented in our previous work,[8] the SD, PNOF2 and PNOF4 approximations satisfied the three positivity conditions,[8] yet when a larger set of molecules is considered, only the SD approximation satisfies the P and G conditions in every system studied in this benchmark. The Q condition is only fulfilled in H_2 by SD, PNOF2, PNOF4, and the other PNOF6 versions (besides PNOF6d).

The P condition checks whether the geminal occupations are non-negative. SD and PNOF6d are the only rDMFAs that fulfil this condition. The PNOF group generates deviations which are an order of magnitude smaller than the other approximations and, in some particular molecules, PNOF2 and PNOF4 do satisfy this condition (see the Suplementary Information). PNOF3 is the rDMFA that produces the largest amount of negative eigenvalues from the PNOF family. K-only functionals produce larger sums of negative eigenvalues than PNOF3, and CA is the approximation that produces the smallest sum in general terms. MBB, BBC2, and CGA approximations produce the largest sum. However, the amount of deviations produced by CA depends on the geometry studied, since, ML and MLSIC produce less deviations in stretched, closed-shell systems. ML, MLSIC, and, in a lesser extent, GU produce consistent deviations independently of the geometry or spin state considered. As expected, the positive character of the eigenvalues of the P

matrix is better preserved at R_{eq}.

The largest averaged sum of negative eigenvalues of the Q matrix is produced by the SD approximation, which conflicts with the fact that the Q condition is met in the hydrogen molecule and the harmonium atom.[8] The CA and ML approximations produce the lowest sum among the K-functionals, with ML being the most consistent for any molecule. SD produces consistent errors for both geometries, but in open-shell molecules larger deviations are committed. As in the results obtained for the P condition, PNOFs generate smaller sums than the K-functionals, and again PNOF3 produces the greatest sum of negative eigenvalues.

The particle-hole probabilities are always positive semidefinite in SD, MBB, CA, CGA, GU, PNOF4, and the three versions of PNOF6, as they were designed to meet the G condition. The rest of rDMFAs, namely BBC2, MLSIC, ML, PNOF2 and PNOF3, present negative eigenvalues in the G matrix, even though the averaged sum is small compared to the sums obtained for the P and Q conditions (excepting the G condition calculated with PNOF2 and PNOF3, for which the sum is slightly larger).

Since the PNOF family of approximations were designed to satisfy as many of these conditions as possible, it is not entirely surprising that these functionals are the ones that produce the smallest sum of negative eigenvalues in the P, Q, and G matrices, being the PNOF6 versions and PNOF4 the rDMFAs that provide the smallest amount of violations. The CA functional presents the lowest sum of negative eigenvalues among the K-functionals, yet larger deviations are obtained at the stretched geometry. ML shows a slightly greater sum of negative eigenvalues, but with a consistent error.

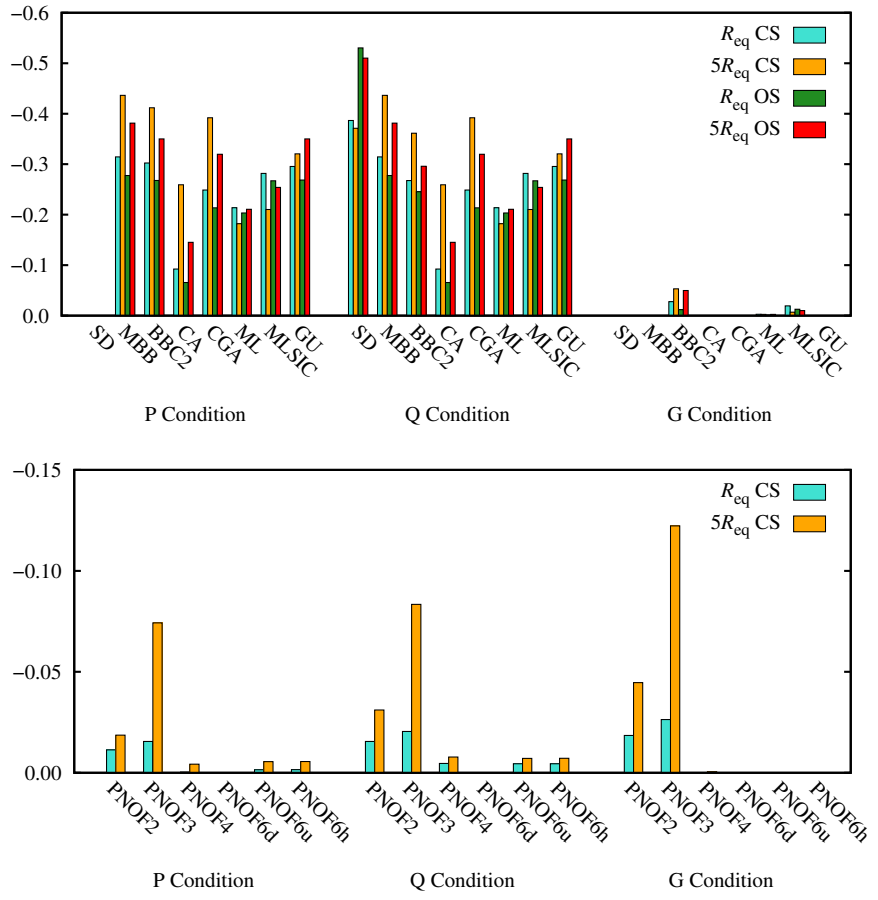


Figure 4: Average of the normalized (by the number of electron pairs) sum of negative eigenvalues from the approximate P, Q and G matrices for closed-shell (CS) and open-shell (OS) molecules.

Antisymmetry

The PNOF family and the SD approximation are designed to produce antisymmetric 2-DMs. The rest of approximations, however, fail to some extent to attend the symmetry condition. Compared to the rest of funtionals, the deviations produced by ML are rather small and stationary in both geometries. The averaged antisymmetry error at $5R_{eq}$ is twice the error at R_{eq} in all 2-DM approximations but MLSIC, where the error is larger at the equilibrium, and ML and GU, which present consistent errors. The magnitude of the averaged error committed for closed-shell molecules is the same for the open-shell molecules set. The MLSIC and GU approximations produce the largest deviations in this test.

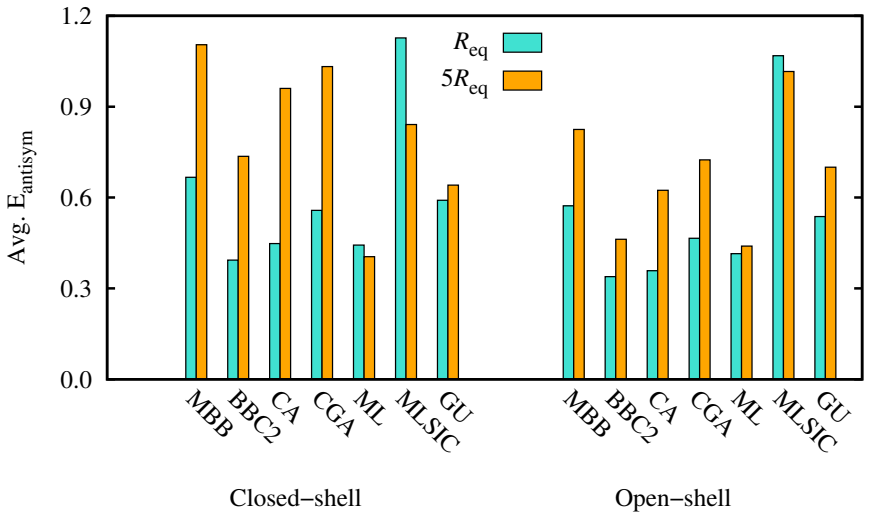


Figure 5: Average of the normalized (by the number of electron pairs) antisymmetry error $E_{antisym}$ (eq. 28).

Performance in physical chemistry properties

Delocalization index

The DI measures the covalent bond order between two fragments and decreases to zero as a bond is stretched and the molecule is dissociated.[59] Despite DI can be negative, they rarely are unless it corresponds to largely separated or unconnected atoms, with an absolute magnitude that is never very large. In the present case, no molecule displays a negative DI at either geometry according to the reference. However, MLSIC and GU mainly, but also PNOF3 and BBC2 to a lesser extent, predict negative DIs. In fact, MLSIC and GU only provided one positive DI for closed-shell molecules at $5R_{eq}$. Even though GU does not predict any negative DI in open-shell molecules, nor BBC2 in the closed-shell ones, these rDMFAs produce the largest errors among the functionals considered. MLSIC and GU, provide worse predictions than their corresponding self-interaction uncorrected versions. BBC2, which is also an improvement with respect to MBB, also presents large deviations with respect its predecessor.

The K-only functionals that do not provide negative DIs present smaller errors at

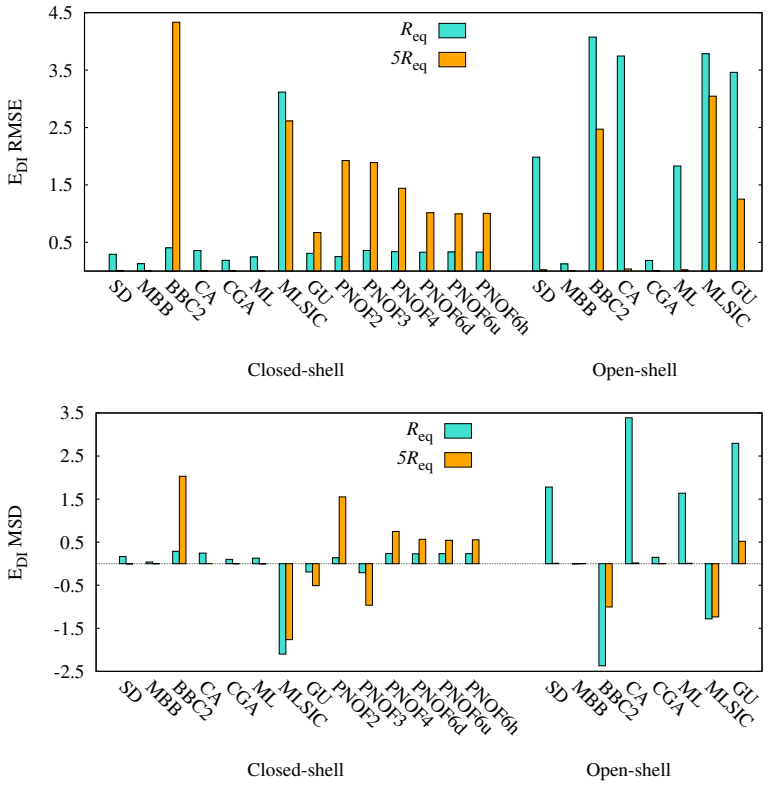


Figure 6: Root mean squared error (RMSE) and mean signed deviation (MSD) for the errors produced in delocalization index, E_{DI} (eq. 29), for closed- and open-shell molecules.

stretched geometries because there is no bond at $5R_{eq}$, and the corresponding DI is zero or practically in most of the diatomic molecules. The K-functionals are able to reproduce the DI in this scenario and, therefore, larger errors are committed at R_{eq}. On the other hand, the PNOF family still predicts a binding value in most of the molecules at $5R_{eq}$, being PNOF2 and PNOF3 the ones that deviate the most. Because the PNOFs predict a bonding interaction still at five times the equilibrium geometry, the K-functionals present a better ability at reproducing the bond order of dissociated molecules.

MBB and CGA produce the smallest deviations among the rDMFAs at both geometries, regardless the spin state of the molecule (see the Supporting Information). On the other hand, whereas the errors produced by SD, CA, ML in closed-shell molecules are considerably small, there is a huge inconsistency in the DI predictions for open-shell molecules.

Generally, all the functionals studied in this work tend to overestimate the covalent bond order of the molecules, excepting functionals that produce negative DIs at large separations. In addition, reproducing the DI in open-shell molecules is clearly a problem for all the approximations.

Mean interelectronic distance $\langle s \rangle$ and variance σ^2

In this section, the average distance between electron pairs $\langle s \rangle$ and its corresponding variance σ^2 are analyzed. The MSD plot presented in Figure 7 indicates that PNOF2 predicts electron pairs to be too close (negative MSD values), and its deviations are larger than the SD approximation. Its description for stretched H_2 is particularly poor, where electron pairs are twice as close as in the reference calculation. PNOF3 provides a corrected description at $5R_{eq}$ with respect to PNOF2, yet H_2 and other molecules are slightly incorrectly described (see the Supporting Infromation). In contrast to the rest of functionals, PNOF3 usually overestimates $\langle s \rangle$ (see the MSD plot). The variances obtained for both PNOF2 and PNOF3 approximations at $5R_{eq}$ are not in agreement with the FCI variance, where the overestimation of $\langle s \rangle$ implies an over-delocalization of electrons (underestimation of σ^2) in PNOF2, and the other way around in PNOF3. Variances and interelectronic distances in stretched molecules are significantly improved in the PNOF4 and PNOF6 versions. Generally, variances are correctly predicted at R_{eq} by both the PNOF family and the K-functionals.

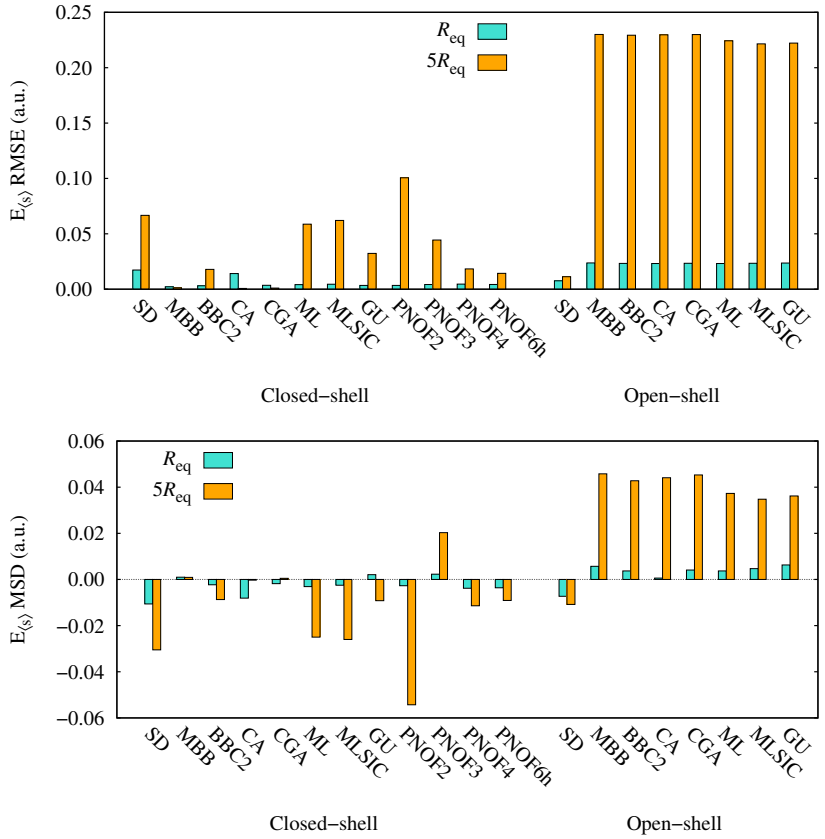


Figure 7: RMSE and MSD for the relative error in the average distance between electron pairs $E_{\langle s \rangle}$ (Eq. 33), for closed- and open-shell molecules. PNOF6u and PNOF6d produce the same deviations as PNOF6h and are thus omitted to avoid redundance.

K-functionals tend to underestimate $\langle s \rangle$, and $5R_{eq}$ geometries are more prone to deviations than equilibrium geometries. ML and MLSIC provide a similar error magnitude than SD, yet their equilibrium predictions are improved with respect to the latter. Along with GU, this group of K-functionals presents larger deviations in the variance only at stretched geometries. The rest of K-functionals (MBB, BBC2, CA ,and CGA) produce similar errors as PNOF4 and the three PNOF6 versions, and all of them provide better predictions for both $\langle s \rangle$ and σ^2 than PNOF2 and PNOF3.

There is a big overestimation of $\langle s \rangle$ in open-shell cases, in which the SD approximation produces more accurate values of $\langle s \rangle$ and σ^2 than any K-functional. Interestingly, whereas the error in $\langle s \rangle$ is larger in open-shell cases, the error in σ^2 is smaller in open-

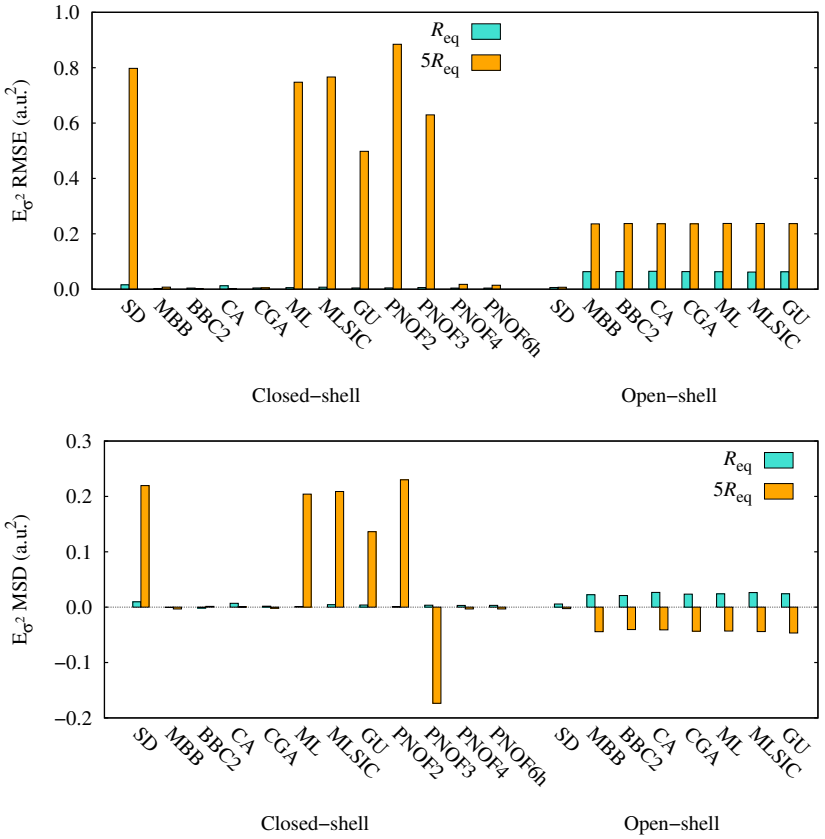


Figure 8: RMSE and MSD for the relative error in the variance of the interelectronic distance E_{σ^2} (Eq. 34), for closed- and open-shell molecules. PNOF6u and PNOF6d produce the same deviations as PNOF6h and are thus omitted to avoid redundance.

shell molecules than in the closed-shell ones.

It is interesting to remark that the largest deviations committed in this test were obtained for the stretched H_2 molecule. PNOF4, PNOF6, MBB, BBC2, CA and CGA are the functionals that best approximate the interelectronic distance value and its variance. Most rDMFAs tend to uncorrectly distribute electron pairs when a molecule is stretched.

Intracule probability densities

Negative values

Because the IPD is a probability distribution of interelectronic distances, its values should be compressed between zero and a positive value to bear a physical significance. Negative probabilities are obtained if the corresponding 2-DM is formed by negative diagonal elements, which are also unphysical. We have calculated the IPD for all the molecules and geometries considered, consulted their profiles, and whether any of the functionals provide any negative values.

In contrast to the IPDs obtained for the harmonium atom,[8] in which most of the functionals presented negative IPD values in the short-range (SR) region, negative IPD values are only predicted in "extremely" long-range (LR) regions in the current set of molecules considered. The region where the negative points are obtained is not that relevant for the electron pair information since the values are obtained around $s_{\text{neg}} \sim 2\langle s \rangle$. That is, the region in which the IPD values are large are found at smaller s compared to the region at which the negative IPD values are obtained. Moreover, the magnitude of the negative values is remarkably small, 10^{-5} being the largest negative value. Therefore, one could consider these values to be residual and not affecting the physical properties of the molecule. This unphysical feature is obtained for both equilibrium and stretched geometries and for most of the molecules considered in this set, mainly predicted by all the K-functionals. However, exceptions are found in stretched H_2, He_2, and Li_2. The amount of negative values in the mid-range (MR) region of He_2 and their magnitude are both small, but BBC2, CA, CGA, GU, MBB, PNOF3, and PNOF6 approximations fail to describe the SR of stretched H_2. This same group of functionals also predicted negative IPD values in the SR region of harmonium atom in the strong-correlation regime.[8] PNOF3 is the functional that provides the worst profile (see Figure 9), PNOF3 providing a negative minimum in the MR of stretched Li_2. Nevertheless, the magnitude of the negative values

obtained for the set of molecules in this current work is two orders smaller than the ones obtained for the harmonium atom.

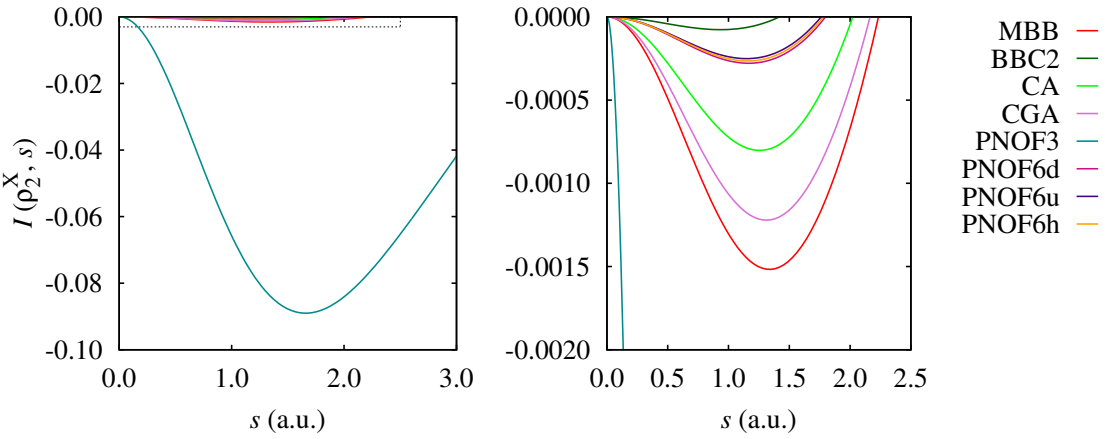


Figure 9: Intracule probability densities (IPDs, eq. 32) of the approximate 2-DM that give unphysical IPD values of H_2 at $5R_{eq}$. The picture on the right contains a zoomed version of the region within the dashed box from the left figure. See the Supporting Information for the complete figure of the IPD.

IPD profiles

In this test, we evaluate the difference between the reference and the approximate IPD ($E_{I(s)}$, see eq. 31). Note that the substraction is the inverse to the one applied in the previous tests, and $E_{I(s)}$ determines the lack of electron correlation with respect to FCI at different range separations. When a method underestimates the correlated motion between electrons, the IPD difference $E_{I(s)}$ is negative in the short-range region (small interelectronic distances s) of the graph, and positive in the long-range region. $E_{I(s)}$ plots for every molecule are gathered in the Supplementary Information.

The $E_{I(s)}$ profiles obtained for open-shell molecules indicate that SD reproduces better the IPD than the K-functionals (which behave similarly to SD only for CH at equilibrium, see Figures S14–S19 in the Supporting Information). In most of the cases, the K-functionals generate analogous and almost undistinguishable IPDs, indicating that the 2-DMs are approximated in a similar way for open-shell cases. Unlike SD, rDMFAs re-

produce the IPD of singlet oxygen instead of the open-shell triplet (compare Figures S11 and S19), which suggests that the approximate 2-DM reproduce closed-shell structures. The SD approximation also describes better the long-range region of the holes, obtaining an almost zero $E_{I(s)}$ for SD, whereas the rest of functionals present larger or lower probabilities, depending on the molecule considered. Another difference in these cases is that the $E_{I(s)}$ in the SD approximation always produces a Coulomb hole shape (*i.e.* negative in the short-range region and positive in the long-range region) for both geometries studied, caused by a general underestimation of V_{ee} and smaller $\langle s \rangle$. Instead, the rest of rDMFAs present different $E_{I(s)}$ shapes according to the molecule examined. For instance, at the equilibrium geometry of CH and OH (Figures S15 and S18), $E_{I(s)}$ is negative at short ranges and positive at larger interelectronic distances; however, rDMFAs produce the inverse profile for the other four open-shell molecules of the set. The magnitude of $E_{I(s)}$ is larger at $5R_{\text{eq}}$, excepting in triplet O_2 (Figures S33–S38). Positive $E_{I(s)}$ values at the long-range region are obtained for stretched CN and OH (Figures S35 and S37), suggesting that K-functionals underestimate electron pair probabilities of largely separated electrons. Conversely, the long-range region of $E_{I(s)}$ is negative for CH, BeH, NO, and triplet O_2 at the stretched geometries. Different $E_{I(s)}$ profiles are generated for the same molecule at two different geometries and, therefore, the behavior of the approximate IPDs generally difficult to predict in open-shell molecules.

Unlike open-shell cases, rDMFAs exhibit different profiles for close-shell molecules (see Figures S1–S13 and S20–S32 for equilibrium and stretched geometries, respectively). This fact puts forward that the current approximations are rather tailored for closed-shell molecules. The magnitude of $E_{I(s)}$ does not change much between both geometries, but in general the error is larger at $5R_{\text{eq}}$. SD produces larger deviations from the reference IPD compared to the rest of the rDMFAs, which is a big improvement with respect to the performance of the rDMFAs in open-shell molecules. PNOF4 and PNOF6 are approximately equivalent through all the systems considered, and $E_{I(s)}$ is negative at the

short-range region and positive at the long-range, indicating a slight underestimation of electron correlation. This profile is obtained at both equilibrium and stretched geometries. These versions of PNOF provide the smallest errors in molecules with a small number of electrons (see Figures S1–S4). PNOF3 and GU present a similar IPD behavior at R_{eq} due to an overestimation of correlation. Because electron pairs are too separated, their $\mathrm{E}_{I(s)}$ is positive in the short-range region and negative in longer ranges. The CA approximation presents irregular, oscillating $\mathrm{E}_{I(s)}$ profiles through all the interelectronic distance axis, and is, in general, the functional that provides the largest deviations. These oscillations are not present when the system analyzed has a small number of electrons, and CA produces an IPD very similar to the SD one (Figures S1–S4). PNOF2, ML, MLSIC, CGA, MBB, and BBC2 show similar $\mathrm{E}_{I(s)}$ trends as PNOF4 and PNOF6, yet they present a posterior maximum at the long-range region of $\mathrm{E}_{I(s)}$ at equilibrium geometries. In general, there is no rDMFA that provides an accurate description of the IPD in a specific range when treating equilibrium geometries (besides PNOF4 and PNOF6 in molecules with a small number of electrons, as already mentioned). The deviations obtained along the interelectronic distance are similar for all rDMFAs, where, in general, the error in the mid-range region is always negative.

The short-range region of the IPD of molecules at $5R_{\mathrm{eq}}$ basically describes the short-range interactions present at R_{eq}, and the profiles at the short-range region of $\mathrm{E}_{I(s)}$ at $5R_{\mathrm{eq}}$ are very similar to the complete $\mathrm{E}_{I(s)}$ picture at equilibrium. On the other and, the long-range region of $\mathrm{E}_{I(s)}$ at the stretched geometries permits to discern whether a method is accurate enough to describe dissociations. The magnitude of $\mathrm{E}_{I(s)}$ in the long-range region is somewhat larger than the one from the short-range region, indicating that the pair-electron distribution between electrons at different atoms is not well described. SD, MBB, CA, CGA, ML, and, to a lesser extent, MLSIC, accurately reproduce the long-range part of $\mathrm{E}_{I(s)}$ of stretched geometries. GU and MLSIC provide quite accurate IPDs through all the interelectronic distances, yet in some cases they overestimate the long-range part

of the IPD. BBC2 and GU reproduce this region correctly only for particular molecules, such as Lif or NH (Figures S24 and S28). The long-range IPD values of PNOF2, PNOF4, and PNOF6 are usually lower than the FCI IPD values (positive $E_{I(s)}$). This behavior is directly related to the results obtained in the DI test, in which the PNOFs predicted a bonding interaction at $5R_{\text{eq}}$. The long-range region of $E_{I(s)}$ in PNOF3 presents different profiles depending on the molecule considered, which indicates that PNOF3 does not present a systematic and predictable error.

We put special attention to the IPD of the stretched geometry of the helium dimer (Figure S21) since it is bonded by weak van der Waals forces (long-ranged).[67] MBB, ML, CGA, CA, SD, PNOF3, and PNOF2 show very small differences with respect to the reference IPD, giving thus a correct description. BBC2, PNOF4, and PNOF6 underestimate the probability of electron pairs at such long interelectronic distances, whereas GU and MLSIC present a small overestimation. These profiles are in agreement with the ones obtained for the DI test (check the DI values of He_2 in the Supporting Information). For further discussion on the ability of the functionals to describe dispersion interactions, see the dispersion test section (*vide infra*).

Energetic analysis

This section includes tests that consider the ability of the rDMFAs to reproduce the electronic repulsion energy. Because in rDMFT the only approximated term of the electronic Hamiltonian is the electron-electron repulsion V_{ee} (eq. 10), we analyze the quality of the predictions of this quantity for each functional in the set of diatomic molecules considered.

Interelectronic repulsion energy V_{ee} in diatomic molecules

Analysis of the electron-electron repulsion energies of the set of diatomic molecules reveals what has already been concluded from the previous IPD and average interelectronic separation $\langle s \rangle$ tests: energetic predictions present more deviations from the reference at stretched geometries (Figure 10). The SD approximation produces more accurate V_{ee} values than the K-functionals in open-shell molecules, which is in agreement with the smaller $\mathrm{E}_{I(s)}$ magnitudes obtained in the IPD test. The error magnitudes in V_{ee} are likewise similar among the K-functionals. SD usually produces a slight overestimation of V_{ee}, whereas K-functionals tend to underestimate it. However, there is neither systematic overestimation nor underestimation of V_{ee}, which depends on the molecule studied. Approximations produce larger V_{ee} errors for closed-shell molecules at stretched geometries, being PNOF2 the one with an RMSE larger than SD, followed by ML, MLSIC, GU, and PNOF3. Instead, the latter group of functionals provides better energetic predictions at R_{eq}, whereas CA and SD are the ones with the largest deviations for both open- and closed-shell molecules. Functionals that commit the largest error are coincident with the ones that were not able to reproduce correctly $\langle s \rangle$ in the corresponding test. MBB, BBC2, CGA, and GU are the K-functionals that perform the best. Unlike K-functionals, PNOF4 and PNOF6 systematically present small overestimations of the electron-electron repulsion energy. Likewise, PNOF2 always presents large overestimations, and PNOF3 is the only approximation of the family that underestimates V_{ee}. Note that the repulsion energy prediction for the latter functional in stretched H_2 is only half the FCI value.

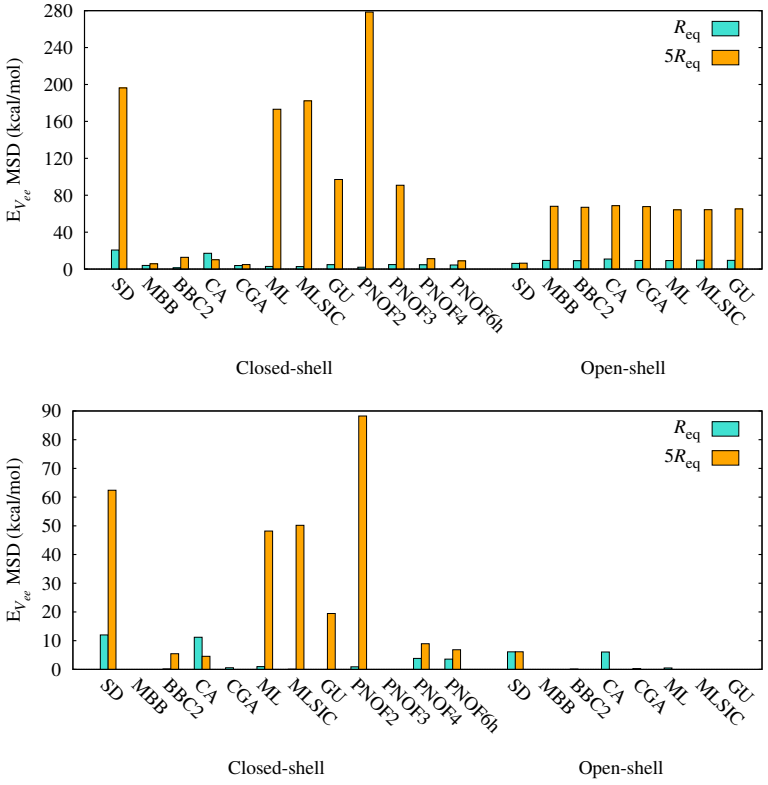


Figure 10: RMSE and MSD for the relative error in the interelectronic repulsion energy $E_{V_{ee}}$ (eq. 36) in atomic units, for closed- and open-shell molecules. PNOF6u and PNOF6d produce the same deviations as PNOF6h and are thus omitted to avoid redundance.

As intuitively expected, the distance between electron pairs $\langle s \rangle$ is directly linked to the repulsive potential V_{ee}. In practically all cases studied, functionals that predict smaller (larger) $\langle s \rangle$ than the FCI reference produce a larger (smaller) electron-electron potential. In agreement with the average electron pair distances test, the MSD plot for the interelectronic repulsion energies reveals the inverse trends seen in Figure 7.

Dissociation energies of dimers

Dissociation energies (D_e) are routinely used in computational chemistry and are a measure of size-consistency. For the molecule to be dissociated, we have considered energies for geometries that are fifty times the equilibrium geometry, $50R_{\text{eq}}$. Since the other energy components are exactly calculated from FCI natural orbital occupancies, by studying D_e we are only considering the error of the interelectronic repulsion potential (eq. 38). The

RMSE and MSD of E_{D_e} are summarized in Figure 11 for all functional approximations.

Some rDMFAs show a severe problem in describing the dissociation potentials, since SD, BBC3, ML, MLSIC, GU, PNOF2, and PNOF3 present a RMSE of the error larger than 100 kcal/mol, where BBC2 has a deviation of almost 400 kcal/mol. This error mostly arises from predictions of V_{ee} at $50R_{\text{eq}}$ geometries in the H_2 molecule, where this group of rDMFAs predict an excessively large repulsive potential at such interelectronic distances (and PNOF3 produces a negative potential). As seen in the V_{ee} prediction test, functionals produce larger errors when considering stretched geometries. Whereas BBC2 was one of the best performing rDMFAs in that test, a too large error in H_2 makes it an unreliable functional for this test. PNOF4 and PNOF6 provide the best dissociation energies among their functional family. Compared to other tests, rDMFAs provide better predictions in the set of open-shell molecules. Again, according to the results obtained from the V_{ee} test and from this test, the interelectronic repulsion is better predicted at equilibrium geometries. This causes large deviations in D_e predictions by some approximations. Therefore, we suggest that future development of rDMFAs should focus on the description of stretched geometries of molecules.

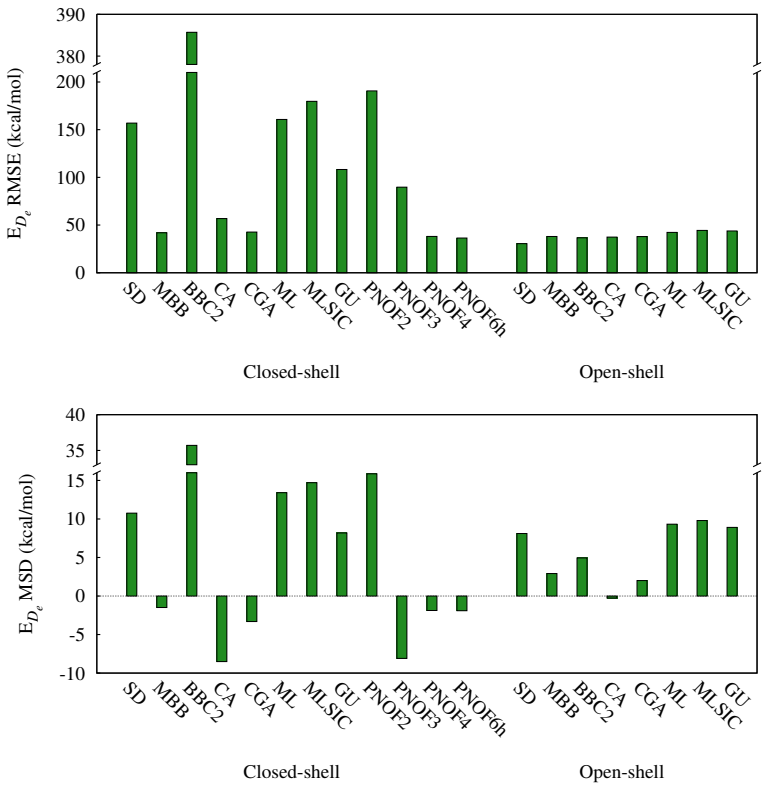


Figure 11: RMSE and MSD for the error in the dissociation interelectronic potentials E_{D_e} (eq. 38), for closed- and open-shell molecules in kcal/mol. PNOF6u and PNOF6d produce the same deviations as PNOF6h and are thus omitted.

$\mathbf{H_4}$ $\mathbf{D_{2h}}$/$\mathbf{D_{4h}}$

The D_{2h}/D_{4h} potential energy surface (PES) of H_4 permits assessing the ability of a functional to describe dynamic and nondynamic correlation effects. The only case in which H_4 is not affected by nondynamic correlation is at the D_{2h} symmetry ($\theta = 70°$) with R = 0.8 Å (the equilibrium geometry[46]). In Figure 12, all the errors reported are represented with respect to this geometry. The PES from going from D_{2h} to D_{4h} but keeping the same distance between hydrogen atoms R (red bars in Figure 12) is better reproduced by the rDMFAs than the other paths analyzed; however, the error committed is of 100 kcal/mol in average. MBB, BBC2, CA, CGA, PNOF2, and PNOF6h produce errors below 65 kcal/mol, but larger than 50 kcal/mol. This shows that rDMFAs can moderately hande nondynamic correlation coming from short-ranged orbital degeneracies. On the other hand, rDMFAs produce errors usually larger than 100 kcal/mol in the other two paths

analyzed (same symmetry as the reference geometry but different R, green bars in Figure 12, and changing both the symmetry and R, blue bars in the figure). The only exceptions are MBB and CGA, and, to a lesser extent, CA and PNOF3 (only in one of the paths). That is, rDMFAs present a problem in describing nondynamic correlation effects arising from the dissociation of a molecule. The errors produced for this situation are equivalent for both symmetries, with the exception of BBC2, which presents a larger deviation at the D_{4h} symmetry. MBB and CGA errors are of 15 kcal/mol in this two paths, being the best performing rDMFAs in this test.

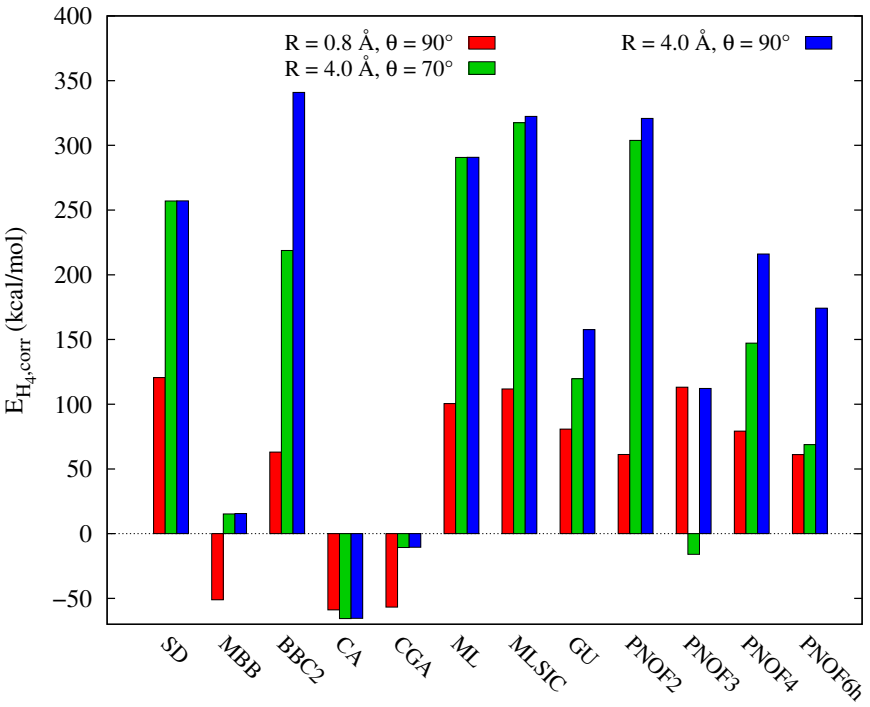


Figure 12: Error of the D_{2h}/D_{4h} potential energy curve of H_4, $E_{H_4,corr}$ (eq. 41), where the repulsion energies are calculated with respect to the interelectronic potential of the ground state geometry, $R = 0.8$ Å and $\theta = 70°$.

A former study carried out in our laboratory on the D_{2h}/D_{4h} potential energy surface of H_4 [46] concluded that PNOF6 produces qualitatively correct energies and orbitals for this system, as PNOF6 includes intrapair and interpair correlation. In contrast to this work, the geometries were optimized using PNOF6. The PNOF6 absolute energies for the equi-

librium geometry from that study produce relative errors of the same order of magnitude as the ones obtained in this work. However, the relative errors do not change drastically when considering other geometries (other angles and hydrogen distances). This suggests that the functional–driven error is larger in PNOF6 than the 1-DM–driven error.

Size-extensivity of the energy: polyhedral H_N

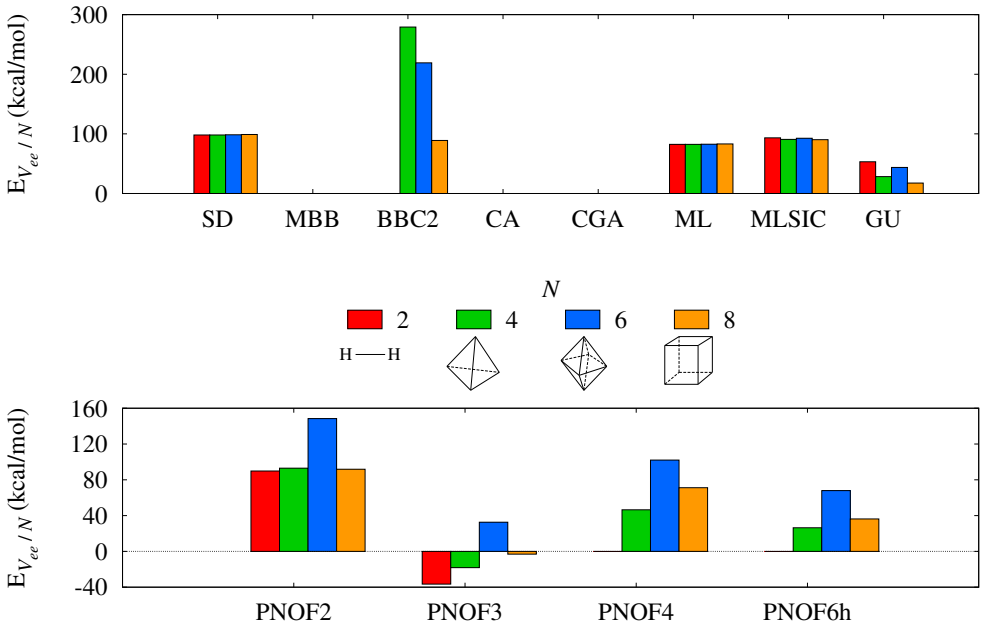


Figure 13: Error in the interelectronic repulsion energy by the number of hydrogen atoms, $E_{V_{ee}/N}$ (eq. 40), for N hydrogen atoms placed at the vertices of an N-vertex polyhedron, in which the H atoms are separated by 10 Å from the geometrical center.

The increase in the number of electrons in the H_N polyhedral molecules implies an increase in the electron-electron repulsion energy. All rDMFAs but BBC2 predict an increase of V_{ee} with the number of atoms; BBC2 predicts larger repulsive potentials when $N = 6$ (see the Supporting Information). Figure 13 illustrates the V_{ee} error per number of atoms N (eq. 40). MBB, CA, and CGA are the only functionals that are able to predict accurate repulsion energies through the H_N molecules considered, with size-extensive results and presenting relative errors below 1 kcal/mol. In contrast, BBC2 shows the largest

deviations, with an unreasonable errors that grow up to 5800 kcal/mol. SD, ML, and ML-SIC present size-extensive energies, but their relative errors are also excessively large. GU and PNOF2 produce smaller errors but they are still non-negligible, and are not the most size extensive rDMFAs studied. PNOF3 is the only approximation that underestimates the interelectronic repulsion, as already discussed in other tests, and predicts negative V_{ee} in $N = 2$. PNOF4 and PNOF6 provide a quality V_{ee} for H_2, but they provide errors of the same order of magnitude as PNOF2 and GU for the rest of H_N molecules. This test enhances the statement of PNOF4 and PNOF6 being good rDMFAs for systems with two electrons, but not the best choice to describe molecules with a larger number of electrons. They do not provide size-extensive energies in this test.

Description of dispersion

The ability of each rDMFA to account for dispersion interactions is analyzed in stretched H_2 and He_2 at different interatomic distances. The dispersion error evaluated, $E_{disp}(R)$ (eq. 42) analyzes whether the rDMFAs account for the IPD condition[65] by multiplying by R^3. In this way, rDMFAs with dependencies higher than R^3 will show an $E_{disp}(R)$ growth to (absolute) large values as R increases. On the other hand, rDMFAs with an exponential or a smaller decay (R^{-4} or lower) will present an $E_{disp}(R)$ that rapidly goes to zero at early values of R.

The black dashed line appearing in both graphs of Figure 14 represents the behavior of the reference. The FCI $E_{disp}(R)$ function is positive and constant when the two fragments are largely separated. Due to a lack of numerical precission, the FCI $E_{disp}(R)$ in H_2 presents a smooth growth with R, caused by a limitation on the angular grids in the present implementation of the RHO2_OPS code, used to perform the IPD integrations.

In both H_2 and He_2, PNOF2, PNOF3, GU, and MLSIC show an $E_{disp}(R)$ that largely

grows or decreases with R, indicating an incorrect description of dispersion interactions caused by a major power dependence with R. Whereas this behavior is not obtained in H_2, PNOF4, PNOF6, and BBC2 also present a larger power dependency with R in He_2. This leads ML, CA, CGA, and MBB the only rDMFAs with a smaller power dependency than R^{-3}.

Another point to be analyzed is the constant value accompanying the $1/R^3$ decay, which is a positive parameter. The positivity of $E_{disp}(R)$ can also indicate the quality of the dispersion interactions treatment. ML, CA, CGA, and MBB present negative E_{disp} values. PNOF2 produces negative $E_{disp}(R)$ values for H_2, but at He_2 $E_{disp}(R)$ are positive, and the contrary behavior is produced by BBC2, PNOF4, and PNOF6. GU, MLSIC, and PNOF3 $E_{disp}(R)$ values are positive in both molecules.

No rDMFA is able to correctly reproduce the universal signature of dispersion interactions in the IPD, and therefore the description of London dispersion forces is not correctly described by any of the rDMFAs. Whereas some functionals produce smaller power dependencies, others bind molecules too strongly. Indeed, in a former study, it was concluded that PNOF2 presents the ability to bind two He atoms to form the van der Waals dimer.[31] However, geometry and, therefore, orbital and occupation number optimizations were carried out in that study, contrarily to this current work. Nevertheless, the PNOF2 results in this test suggest that the higher power dependency with R gives PNOF2 the ability to bind He_2. We believe that such behavior could be corrected if the dispersion condition for the IPD[65] is considered in their development.

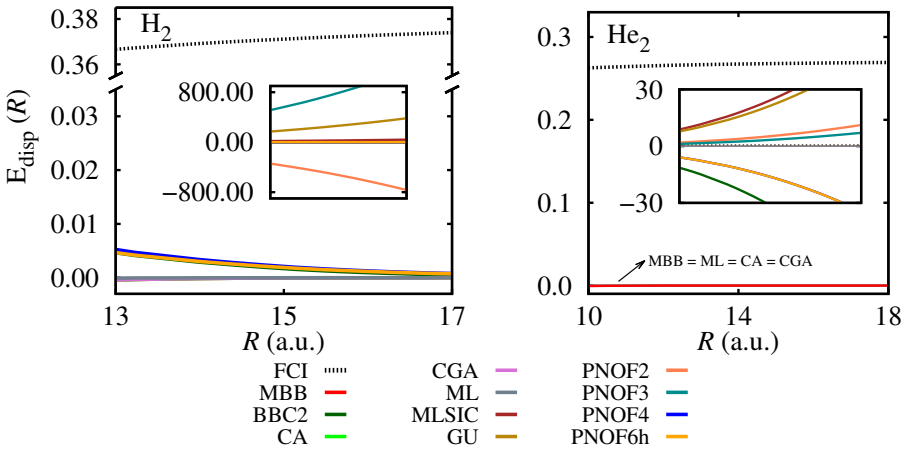


Figure 14: $E_{disp}(R)$ profiles (Eq. 42) in (left) the hydrogen molecule and (right) the helium van der Waals dimer.

Conclusions

A set of thirteen approximations to the 2-DM have been analyzed by considering their physical background in their design, and their performance in predicting chemical properties. No rDMFA has succeeded in all the tests considered in this work, yet their strengths and flaws have been identified and exposed. The MBB, CA, and CGA approximations have provided really good predictions for the energies in the tests carried out in this study. These approximated 2-DMs bear the correct trace, yet they only fulfill the G-condition. They reproduced quite accurately the DIs and predicted quite accurate interelectronic distances $\langle s \rangle$. However, the IPD profiles obtained are quite inaccurate compared to the exact, FCI one, giving very irregular profiles in all the dimers studied in this work. The MBB IPD is, however, one of the best rDMFAs to reproduce the reference IPD. Their performance in open-shell molecules is not that unsatisfactory, since CGA and MBB have provided the smallest deviations in the DI and V_{ee} predictions. However, as the rest of K-functionals, their IPDs and, consequently, the expected value of the interelectronic distance $\langle s \rangle$ are incorrect. In former benchmarks, it has been pointed out that CA and CGA give correct energies but for the wrong reason.[9] In general, MBB has provided quite better results than CA and CGA, but the antisymmetry of the approximate 2-DM is not better

preserved than in CA or CGA. In general, CA, CGA, and MBB approximations tend to provide better predictions than the SD 2-DM, a functional that was not designed to be a rDMFA.

GU and BBC2 show significative worse predictions than their predecessor, MBB. The sums of negative eigenvalues of the P, Q, and G matrices of BBC2 are slightly smaller with respect to MBB, yet the G-condition is not satisfied in BBC2. Instead, the GU 2-DM commits approximately the same violations as MBB, and is G-representable. Their 2-DM elements differ significantly from the FCI ones, and it results in, for instance, bad approximations of the DI, even providing negative DI values that do not correspond with the studied molecules. However, the IPD and its derived properties ($\langle s \rangle$, σ^2, and V_{ee}) are quite accurately predicted by BBC2. BBC2 was explicitly designed to predict energies, but only for non-strongly correlated cases, as it has been seen in the predictions of stretched geometries, the D_e test, and in the H_4 D_{2h}/D_{4h} potential energy surface (see as well the results obtained in the harmonium atom). Instead, the IPD of the GU 2-DM does not improve with respect to MBB, and its energetics, while they are not the worst predictions, they are not the best ones either. Better predictions are also obtained in systems that are not dominated by nondynamic correlation. Instead, BBC2 is able to provide a quite correct long-range behavior that permits describing dispersion interactions quite properly. This result is not obtained in MBB, and a worse behavior is seen in GU.

The number of violations committed by the PNOFs in the positivity matrices is smaller than in the K-functionals group, as the PNOF design imposes these conditions in their structure. PNOF2 and PNOF3 present a small number of violations, and fewer deviations are committed in PNOF4 and PNOF6. PNOF2 and PNOF3 show important deviations in their IPDs, and stretched molecules seem to be problematic to describe. However, the electron pair distribution seen in the IPDs reveals that the PNOFs have improved through the versions. However, PNOF3 provides some unphysical and incorrect descriptions in

most of the molecules examined, being PNOF2 a more appropriate functional than its successor. The latest versions of PNOF considered in this work provide reasonable IPD profiles for dissociating geometries, really similar to the FCI results in cases where electron correlation is small, and reasonably good predictions for V_{ee}. However, their DI predictions fail to describe a dissociated bond order, as well as other energy predictions in cases where nondynamic correlation plays an important role. Instead, they provide a relatively proper treatment of dispersion interactions, in which their older versions, PNOF2 and PNOF3, failed considerably in describing the long-range asymptotic decay.

Table 5 summarizes the results of the tests performed through this article, and gathers the best performing rDMFAs. In general, the rDMFAs that produce the smallest deviations in all the tests are MBB, CA, CGA, PNOF4, and PNOF6 (not necessarily in this order). Newer PNOF versions provide better predictions, yet the oldest K-functionals are the best performing. It does not seem that the self-interaction corrections provide any advantage on the functionals either.

The foundations on rDMFT are almost 50 years old,[1] yet the theoretical research of rDMFT has increased in the last years. Several functionals have been presented through the years and large improvements are obtained. However, there is still the need to improve the performance of 2-DM approximations in the description of open-shell molecules, as well as stretched bonds, as it has been seen through the tests presented in this work. In general, they present difficulties in predicting the IPD of open-shell molecules, where SD provides a better prediction than the rest of K-functionals. The treatment of open-shell molecules is currently being assessed, with promising results.[68–70]

Table 5: Summary of the best-performing rDMFAs according to the results and errors obtained in the tests studied throughout this work. [a)] For open-shell molecules. [b)] At stretched geometries, regions close to interfragment distances, $s = R$. [c)] For systems with small number of electrons. [d)] These functionals provide the smallest deviations, but the RMSE of the error is of 50 kcal/mol. [e)] CGA and MBB are not able to describe short-range correlation arising from orbital degeneracies

Property	Best-performing rDMFA
Trace	MBB, BBC2, CA, CGA, PNOF2, PNOF4
Diagonal elements of the 2-DM	PNOF6, PNOF4, PNOF2
Elements of the 2-DM	PNOF6, PNOF4, SD
N-representability	PNOF family, CA, SD
Antisymmetry	SD, PNOF family
Delocalization index	MBB, CGA, CA, ML, SD
Intracule probability densities	MBB, ML, MLSIC, SD, [a] *K*-functionals, [b] PNOF4, [c] PNOF6 [c]
Average interelectronic distance and variance	MBB, CGA, CA, BBC2
Interelectronic potential	MBB, CGA, CA, PNOF4, PNOF6
Dissociation energies	MBB,[d] CGA, [d] PNOF4, [d] PNOF6, [d] CA [d]
Nondynamic correlation-including	MBB, [e] CGA [e]
Size-extensivity	MBB, CA, CGA
Description of dispersion interactions	none

References

(1) Gilbert, T. Hohenberg-Kohn theorem for nonlocal external potentials. *Phys. Rev. B* **1975**, *12*, 2111.

(2) Piris, M.; Ugalde, J. M. Perspective on natural orbital functional theory. *Int. J. Quantum Chem.* **2014**, *114*, 1169–1175.

(3) Pernal, K.; Giesbertz, K. J. H. In *Density-Functional Methods for Excited States*;

Ferré, N., Filatov, M., Huix-Rotllant, M., Eds.; Springer International Publishing, 2016; pp 125–183.

(4) Herbert, J. M.; Harriman, J. E. Comparison of two-electron densities reconstructed from one-electron density matrices. *Int. J. Quantum Chem.* **2002**, *90*, 355–369.

(5) Herbert, J. M.; Harriman, J. E. *N*-representability and variational stability in natural orbital functional theory. *J. Chem. Phys.* **2003**, *118*, 10835–10846.

(6) Herbert, J. M.; Harriman, J. E. Self-interaction in natural orbital functional theory. *Chem. Phys. Lett.* **2003**, *382*, 142–149.

(7) Cioslowski, J.; Piris, M.; Matito, E. Robust validation of approximate 1-matrix functionals with few-electron harmonium atoms. *J. Chem. Phys.* **2015**, *143*, 214101.

(8) Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Via-Nadal, M.; Piris, M.; Matito, E. Comprehensive benchmarking of density matrix functional approximations. *Phys. Chem. Chem. Phys.* **2017**, *19*, 24029–24041.

(9) Mitxelena, I.; Piris, M.; Rodríguez-Mayorga, M. On the performance of natural orbital functional approximations in the Hubbard model. *J. Phys.: Condens. Matter* **2017**, *29*, 425602.

(10) McWeeny, R. Some recent advances in density matrix theory. *Rev. Mod. Phys.* **1960**, *32*, 335–369.

(11) Coleman, A. J.; Yukalov, V. I. *Reduced density matrices: Coulson's challenge*; Springer Verlag: Berlin, 2000; Vol. 72.

(12) Mazziotti, D. A. Structure of fermionic density matrices: complete *N*-representability conditions. *Phys. Rev. Lett.* **2012**, *108*, 263002.

(13) Coleman, A. J. Structure of Fermion Density Matrices. *Rev. Mod. Phys.* **1963**, *35*, 668–687.

(14) Davidson, E. R. *Reduced density matrices in quantum chemistry*; Academic Press, Inc.: New York, 1976; Vol. 6.

(15) Erdahl, R. M. Representability. *Int. J. Quantum Chem.* **1978**, *13*, 697–718.

(16) Weinhold, F.; Wilson Jr, E. B. Reduced Density Matrices of Atoms and Molecules. II. On the N-Representability Problem. *J. Chem. Phys.* **1967**, *47*, 2298–2311.

(17) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, Ltd: Chichester, 2000.

(18) Mazziotti, D. A.; Erdahl, R. M. Uncertainty relations and reduced density matrices: Mapping many-body quantum mechanics onto four particles. *Phys. Rev. A* **2001**, *63*, 042113.

(19) Gidofalvi, G.; Mazziotti, D. A. Molecular properties from variational reduced-density-matrix theory with three-particle *N*-representability conditions. *J. Chem. Phys.* **2007**, *126*, 024105.

(20) Piris, M. A generalized self-consistent-field procedure in the improved BCS theory. *J. Math. Chem.* **1999**, *25*, 47–54.

(21) Buijse, M. A. Thesis: Electron Correlation. Fermi and Coulomb holes, dynamical and nondynamical correlation. Ph.D. thesis, Vrije Universiteit, Amsterdam, The Netherlands, 1991.

(22) Buijse, M. A.; Baerends, E. J. An approximate exchange-correlation hole density as a functional of the natural orbitals. *Mol. Phys.* **2002**, *100*, 401–421.

(23) Löwdin, P.-O. 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. *Phys. Rev.* **1955**, *97*, 1474–1489.

(24) Müller, A. M. K. Explicit approximate expression between reduced two- and one-particle density matrices. *Phys. Lett.* **1984**, *105A*, 446–452.

(25) Goedecker, S.; Umrigar, C. J. Natural Orbital Functional for the Many-Electron Problem. *Phys. Rev. Lett.* **1998**, *81*, 866–869.

(26) Gritsenko, O.; Pernal, K.; Baerends, E. J. An improved density matrix functional by physically motivated repulsive corrections. *J. Chem. Phys.* **2005**, *122*, 204102.

(27) Csányi, G.; Arias, T. A. Tensor product expansions for correlation in quantum many-body systems. *Phys. Rev. B* **2000**, *61*, 7348.

(28) Csányi, G.; Goedecker, S.; Arias, T. A. Improved tensor-product expansions for the two-particle density matrix. *Phys. Rev. A* **2002**, *65*, 032510.

(29) Marques, M. A. L.; Lathiotakis, N. N. Empirical functionals for reduced-density-matrix-functional theory. *Phys. Rev. A* **2008**, *77*, 032509.

(30) Piris, M. A New Approach for the Two-Electron Cumulant in Natural Orbital Functional Theory. *Int. J. Quantum Chem.* **2006**, *106*, 1093–1104.

(31) Piris, M.; Lopez, X.; Ugalde, J. M. Dispersion interactions within the Piris natural orbital functional theory: The helium dimer. *J. Chem. Phys.* **2007**, *126*, 214103.

(32) Piris, M.; Matxain, J. M.; Lopez, X.; Ugalde, J. M. Communications: Accurate description of atoms and molecules by natural orbital functional theory. *J. Chem. Phys.* **2010**, *132*, 031103.

(33) Piris, M.; Matxain, J. M.; Lopez, X.; Ugalde, J. M. Communication: The role of the positivity N-representability conditions in natural orbital functional theory. *J. Chem. Phys.* **2010**, *133*, 111101.

(34) Piris, M.; Lopez, X.; Ruipérez, F.; Matxain, J. M.; Ugalde, J. M. A natural orbital functional for multiconfigurational states. *J. Chem. Phys.* **2011**, *134*, 164102.

(35) Piris, M. Interacting pairs in natural orbital functional theory. *J. Chem. Phys.* **2014**, *141*, 044107.

(36) Piris, M. Global Method For The Electron Correlation. *Phys. Rev. Lett.* **2017**, *119*, 063002.

(37) Piris, M.; Matxain, J.; Lopez, X. The intrapair electron correlation in natural orbital functional theory. *J. Chem. Phys.* **2013**, *139*, 234109–234109.

(38) Piris, M.; March, N. H. Low-Lying Isomers of Free-Space Halogen Clusters with Tetrahedral and Octahedral Symmetry in Relation to Stable Molecules Such as SF_6. *J. Phys. Chem. A* **2015**, *119*, 10190–10194.

(39) Sim, E.; Song, S.; Burke, K. Quantifying density errors in DFT. *J. Phys. Chem. Lett.* **2018**, *9*, 6385–6392.

(40) Noga, J.; Kutzelnigg, W.; Klopper, W. CC-R12, a correlation cusp corrected coupled-cluster method with a pilot application to the Be_2 potential curve. *Chem. Phys. Lett.* **1992**, *199*, 497–504.

(41) Schmidt, M. W.; Ivanic, J.; Ruedenberg, K. Electronic structure analysis of the ground-state potential energy curve of Be2. *J. Phys. Chem. A* **2010**, *114*, 8687–8696.

(42) El Khatib, M.; Bendazzoli, G. L.; Evangelisti, S.; Helal, W.; Leininger, T.; Tenti, L.; Angeli, C. Beryllium dimer: A bond based on non-dynamical correlation. *J. Phys. Chem. A* **2014**, *118*, 6664–6673.

(43) Juhász, T.; Mazziotti, D. A. The cumulant two-particle reduced density matrix as a measure of electron correlation and entanglement. *J. Chem. Phys.* **2006**, *125*, 174105.

(44) Via-Nadal, M.; Rodríguez Mayorga, M.; Ramos-Cordoba, E.; Matito, E. Singling Out Dynamic and Nondynamic Correlation. *J. Phys. Chem. Lett.* **2019**, *10*, 4032–4037.

(45) Kowalski, K.; Jankowski, K. Towards Complete Solutions to Systems of Nonlinear Equations of Many-Electron Theories. *Phys. Rev. Lett.* **1998**, *81*, 1195–1198.

(46) Ramos-Cordoba, E.; Lopez, X.; Piris, M.; Matito, E. H_4: A challenging system for natural orbital functional approximations. *J. Chem. Phys.* **2015**, *143*, 164112.

(47) Ramos-Cordoba, E.; Salvador, P.; Matito, E. Separation of dynamic and nondynamic correlation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24015–24023.

(48) NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101. Release 20, August 2019, Editor: Russell D. Johnson III. http://cccbdb.nist.gov/,.

(49) Matito, E.; Feixas, F. DMn program. 2009; Universitat de Girona (Spain) and Uniwersytet Szczeciński (Poland).

(50) Michel Caffarel, E. G., Thomas Applencourt; Scemama, A. *Recent Progress in Quantum Monte Carlo*; ACS Publications, 2016; Chapter 2, pp 15–46.

(51) Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A mountaineering strategy to excited states: highly accurate reference energies and benchmarks. *J. Chem. Theory Comput.* **2018**, *47*, 4360–4379.

(52) Dash, M.; Moroni, S.; Scemama, A.; Filippi, C. Perturbatively selected configuration-interaction wave functions for efficient geometry optimization in quantum Monte Carlo. *J. Chem. Theory Comput.* **2018**, *14*, 4176–4182.

(53) Rodríguez-Mayorga, M. RHO-OPS: Density Operations. 2015; Institut de Química Computacional i Catàlisi (IQCC), Universitat of Girona, Catalonia, Spain.

(54) Salvador, P.; Ramos-Cordoba, E. APOST-3D program. 2012; Universitat de Girona (Spain).

(55) Rodríguez-Mayorga, M. RHO2-OPS: 2-DM Operations. GNU GPLv3 public license. 2016; Institut de Química Computacional i Catàlisi (IQCC), Universitat of Girona, Catalonia, Spain.

(56) Cioslowski, J.; Liu, G. Fast evaluation of electron intracule and extracule densities on large grids of points. *J. Chem. Phys.* **1996**, *105*, 4151–4158.

(57) Bader, R. F. W.; Stephens, M. E. Fluctuation and correlation of electrons in molecular systems. *Chem. Phys. Lett.* **1974**, *26*, 445–449.

(58) Fradera, X.; Austen, M. A.; Bader, R. F. W. The Lewis Model and Beyond. *J. Phys. Chem. A* **1999**, *103*, 304–314.

(59) Matito, E.; Solà, M.; Salvador, P.; Duran, M. Electron sharing indexes at the correlated level. Application to aromaticity calculations. *Faraday Discuss.* **2007**, *135*, 325–345.

(60) Eddington, A. S. *Fundamental theory*; Cambridge University Press: Cambridge, 1946.

(61) Thakkar, A. J. In *Density Matrices and Density Functionals*; Erdahl, R. M., Smith Jr., V. H., Eds.; Springer, Doerdrecht, 1987; pp 553–581.

(62) Coleman, A. J. Density matrices in the quantum theory of matter: Energy, intracules and extracules. *Int. J. Quantum Chem.* **1967**, *1*, 457–464.

(63) Boyd, R.; Ugalde, J. Computational Chemistry part A. Ph.D. thesis, ed S Fraga (Amsterdam: Elsevier) p 273–299, 1992.

(64) Coulson, C. A. Present state of molecular structure calculations. *Rev. Mod. Phys.* **1960**, *32*, 170–177.

(65) Via-Nadal, M.; Rodríguez-Mayorga, M.; Matito, E. A Salient Signature of van der Waals Interactions. *Phys. Rev. A* **2017**, *96*, 050501.

(66) Goedecker, S.; Umrigar, C. J. *Many-Electron Densities and Reduced Density Matrices*; Springer: New York, 2000; Chapter 8, pp 165–181.

(67) Stone, A. J. *The theory of intermolecular forces*, 2nd ed.; Oxford University Press, 2013.

(68) Leiva, P.; Piris, M. Description of high-spin restricted open-shell molecules with the Piris natural orbital functional. *Int. J. Quantum Chem.* **2007**, *107*, 1–11.

(69) Theophilou, I.; Lathiotakis, N. N.; Helbig, N. Conditions for Describing Triplet States in Reduced Density Matrix Functional Theory. *J. Chem. Theory Comput.* **2016**, *12*, 2668–2678.

(70) Piris, M. Natural orbital functional for multiplets. *Phys. Rev. A* **2019**, *100*, 032508.

Supplementary Information: A density matrix functional approximations benchmark for atoms and molecules

Mireia Via-Nadal,[†,‡] Eloy Ramos-Cordoba,[†,‡] Pierre François Loos,[¶] and Eduard Matito[*,†,§]

Donostia International Physics Center (DIPC), 20018 Donostia, Euskadi, Spain, Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Euskadi, Spain, Laboratoire de Chimie et Physique Quantiques (UMR 5626), Université de Toulouse, CNRS, UPS, France, and IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Euskadi, Spain

E-mail: ematito@gmail.com

[*]To whom correspondence should be addressed
[†]Donostia International Physics Center (DIPC), 20018 Donostia, Euskadi, Spain
[‡]Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Euskadi, Spain
[¶]Laboratoire de Chimie et Physique Quantiques (UMR 5626), Université de Toulouse, CNRS, UPS, France
[§]IKERBASQUE, Basque Foundation for Science, 48013 Bilbao, Euskadi, Spain

1 Radial intracule probabilities

1.1 Equilibrium geometries R_{eq}

1.1.1 Closed-shell molecules

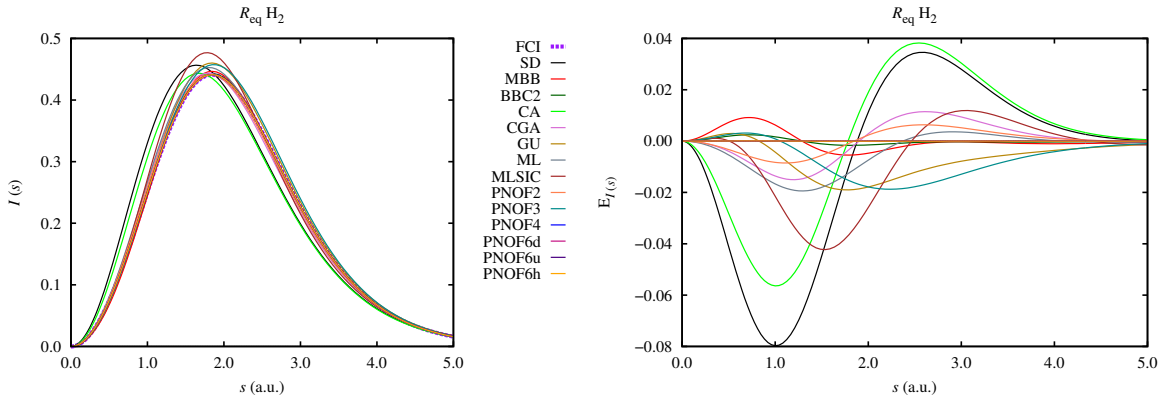


Figure S1: H_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. Left) The exact IPD (FCI), $I(\rho_2, s)$ and the IPDs obtained from the approximate 2-DM, $I(\rho_2^X, s)$, where X stands for a RDMFA. Right) The error in the IPD, $E_{I(s)} = I(\rho_2, s) - I(\rho_2^X, s)$ (eq. 35 in the manuscript).

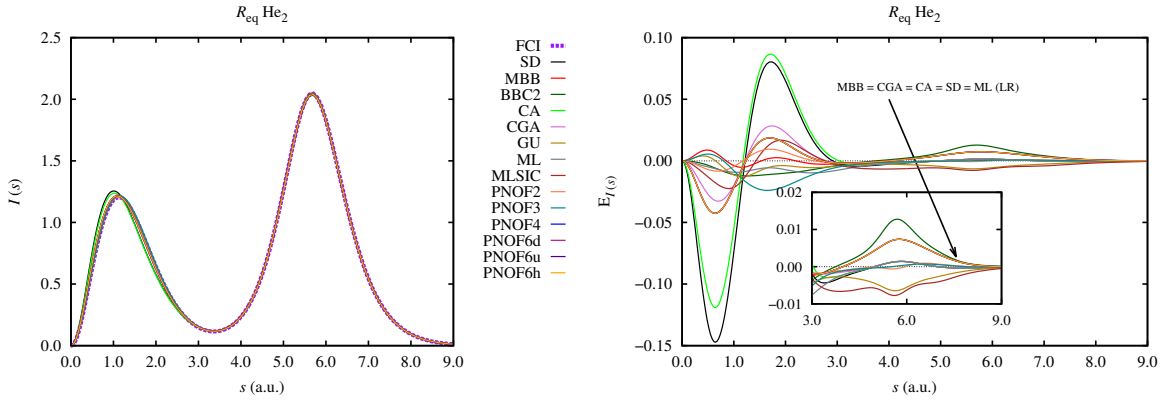


Figure S2: He_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

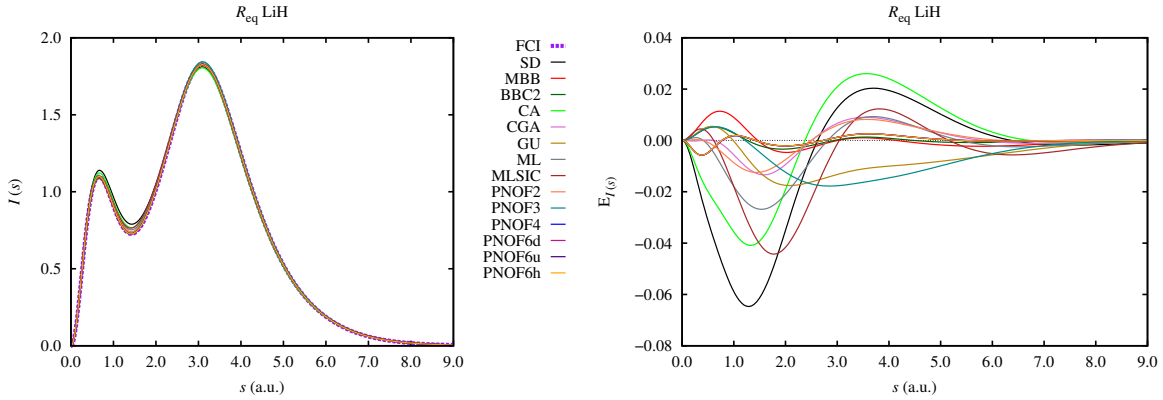


Figure S3: LiH intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

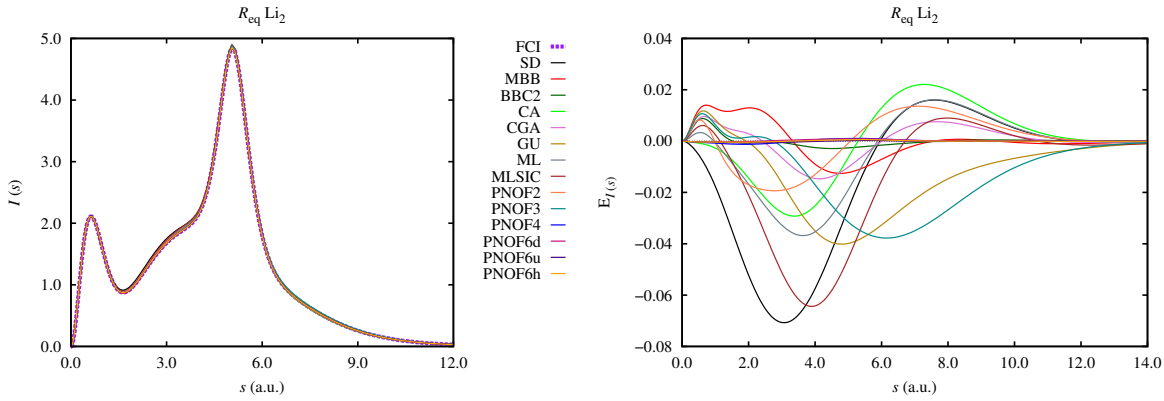


Figure S4: Li_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

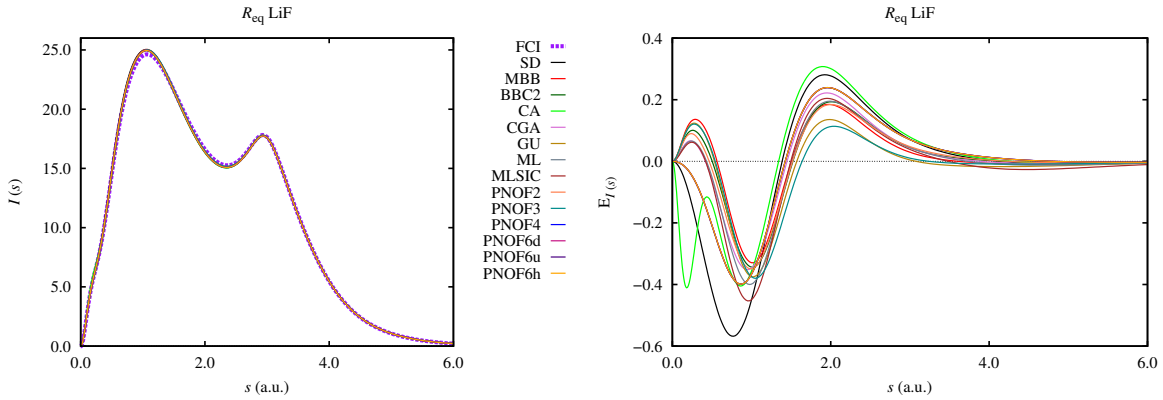


Figure S5: LiF intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

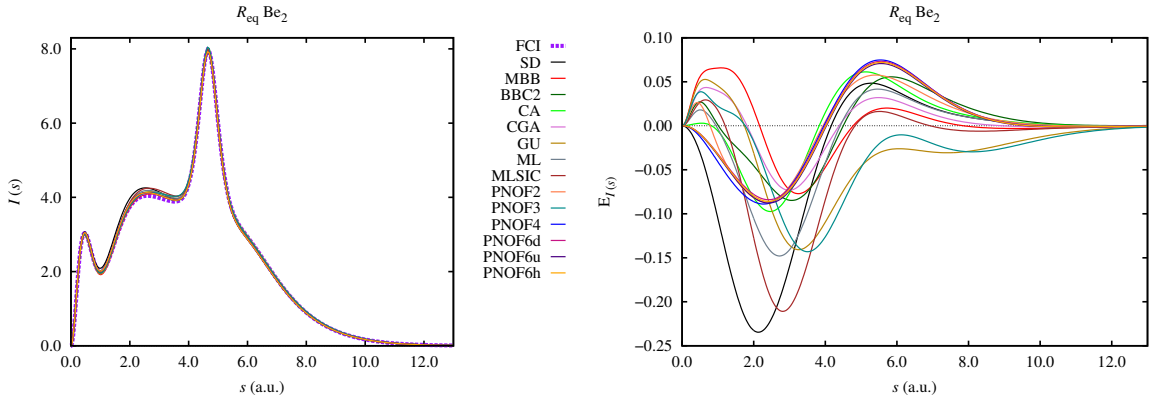


Figure S6: Be_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

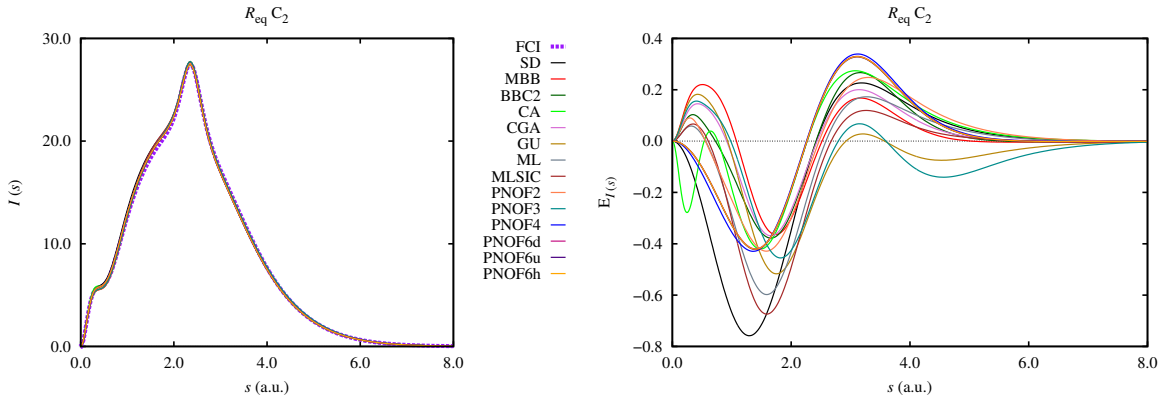


Figure S7: C_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

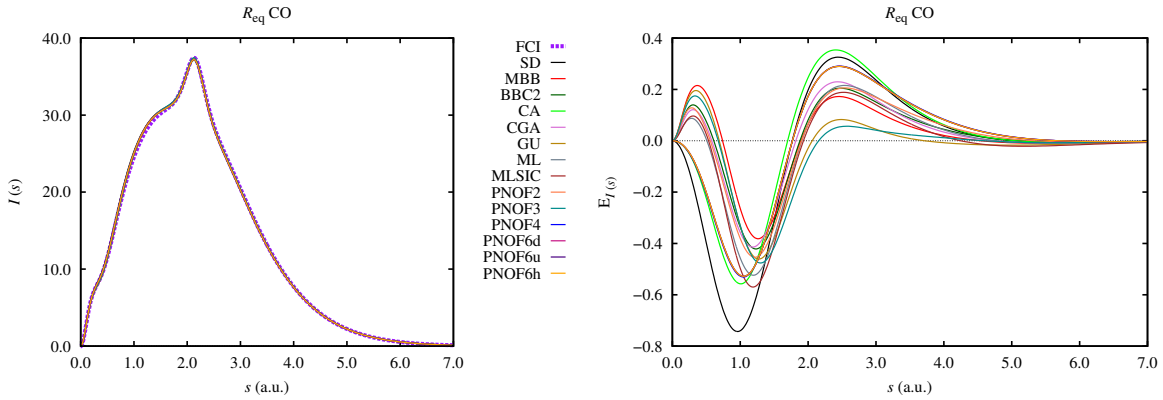


Figure S8: CO intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

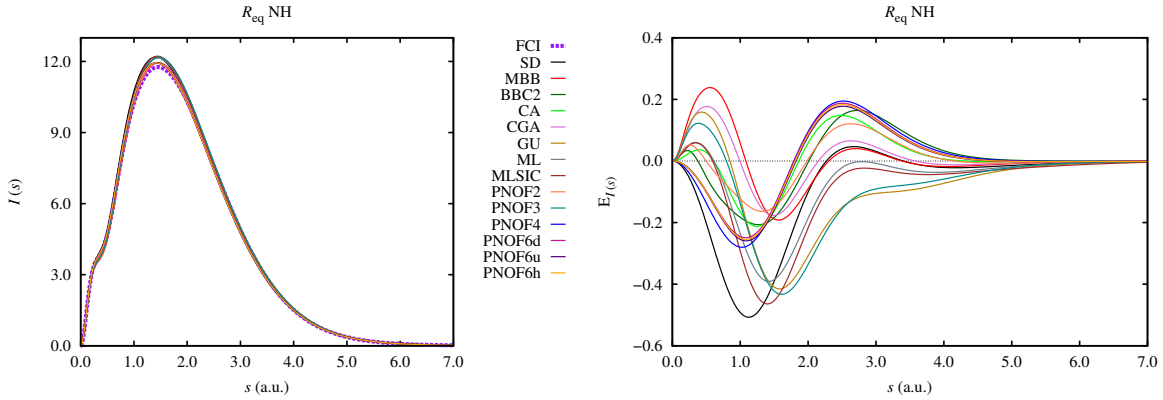


Figure S9: NH intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

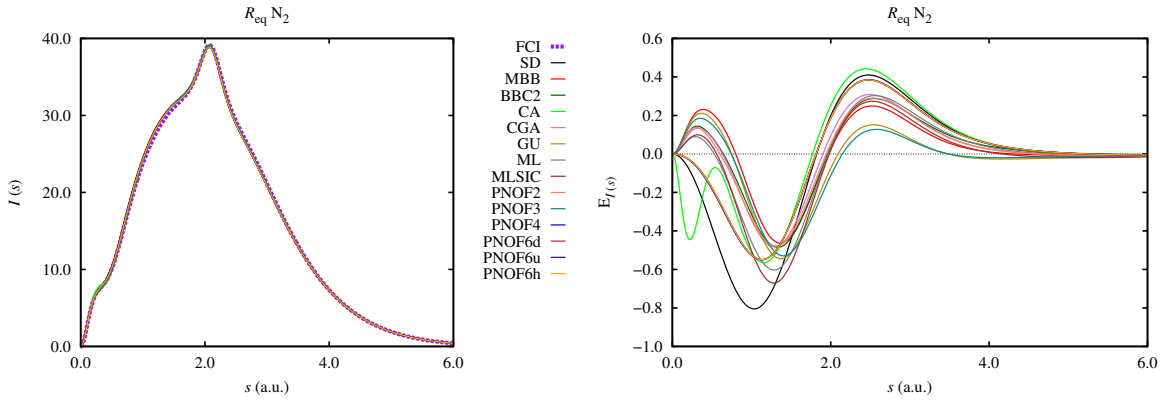


Figure S10: N_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

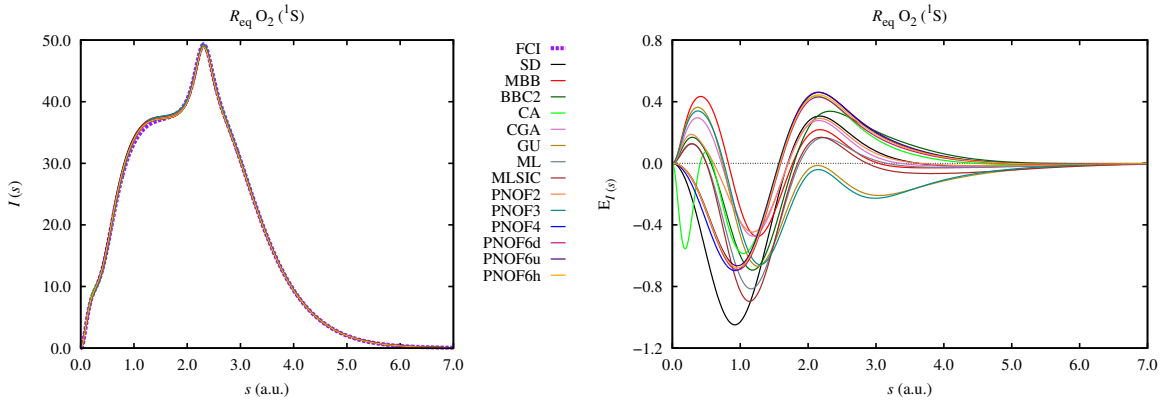


Figure S11: Singlet O_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

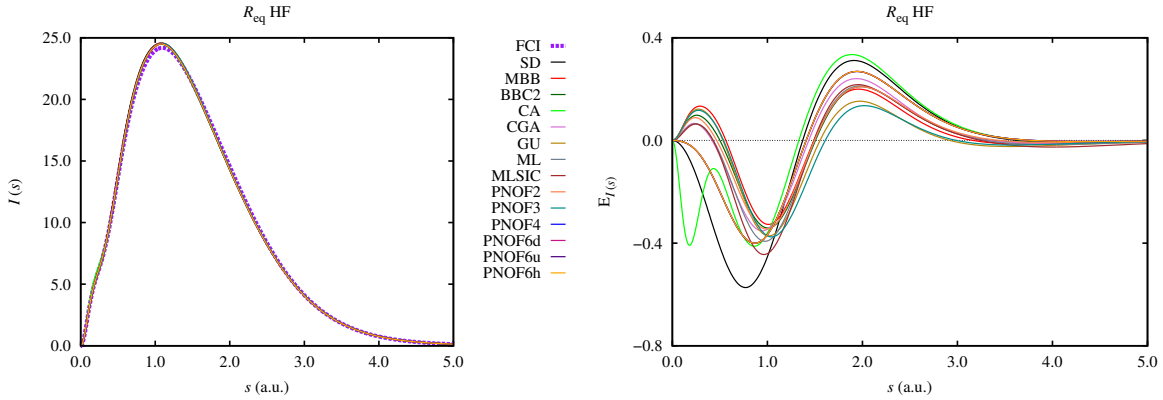


Figure S12: HF Intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

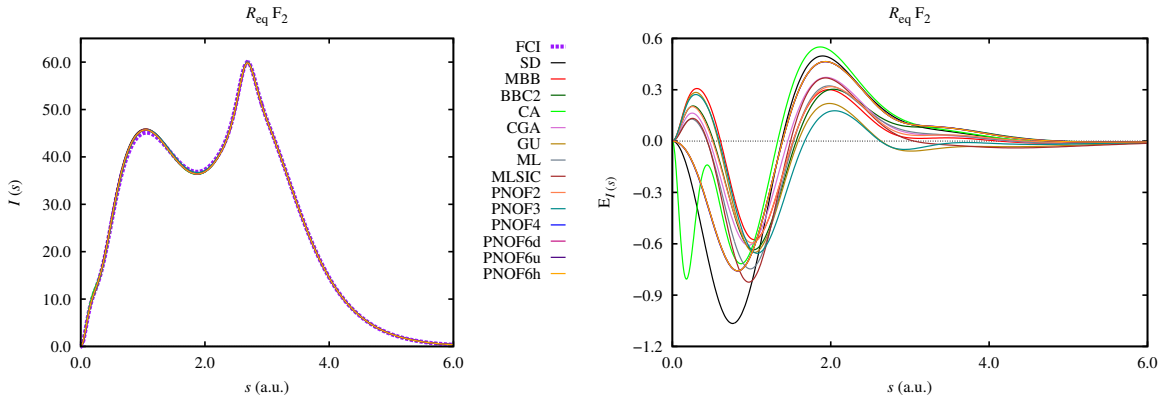


Figure S13: F_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

1.1.2 Open-shell molecules

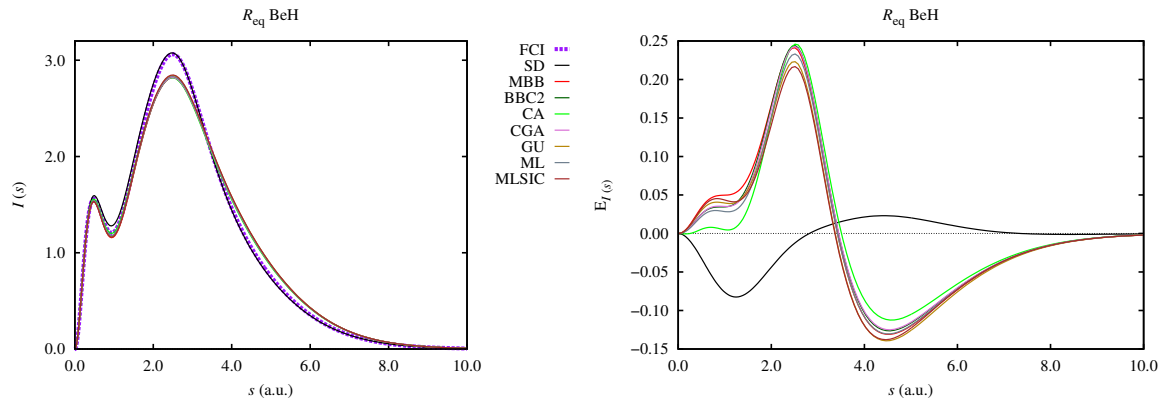


Figure S14: BeH intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

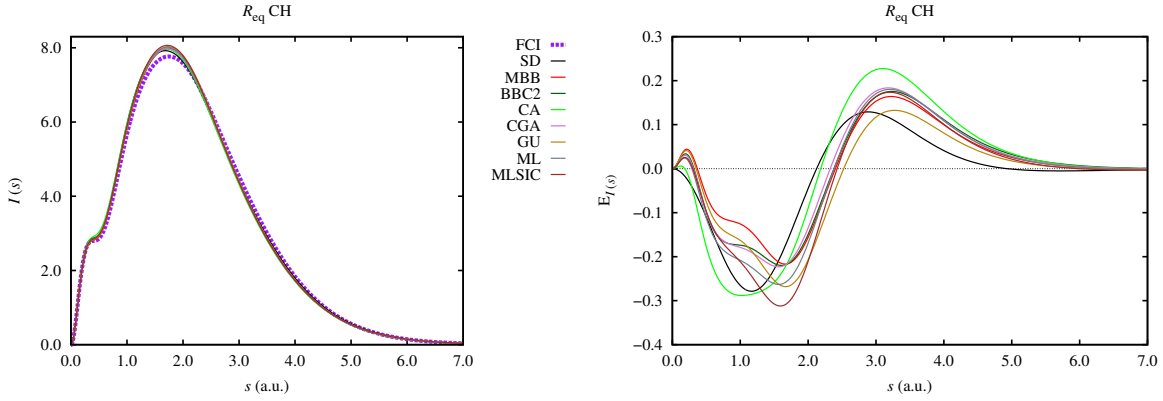


Figure S15: CH intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

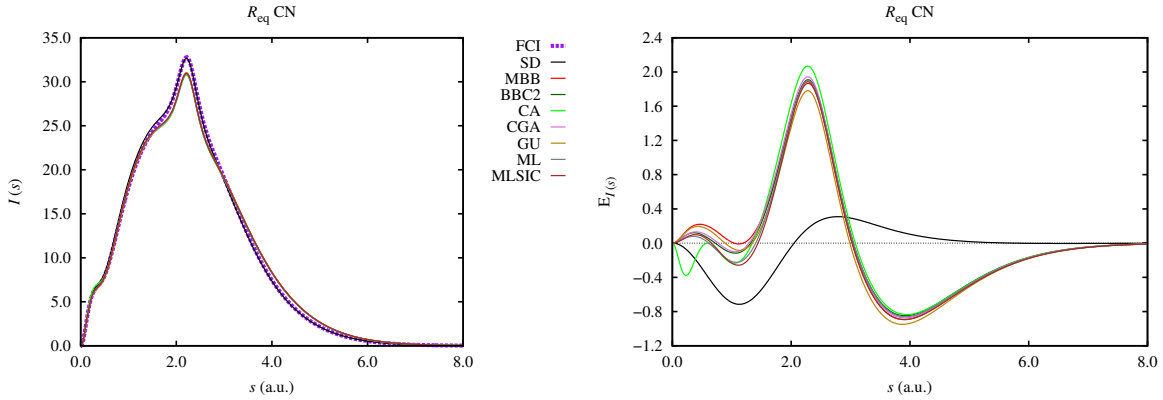


Figure S16: CN intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

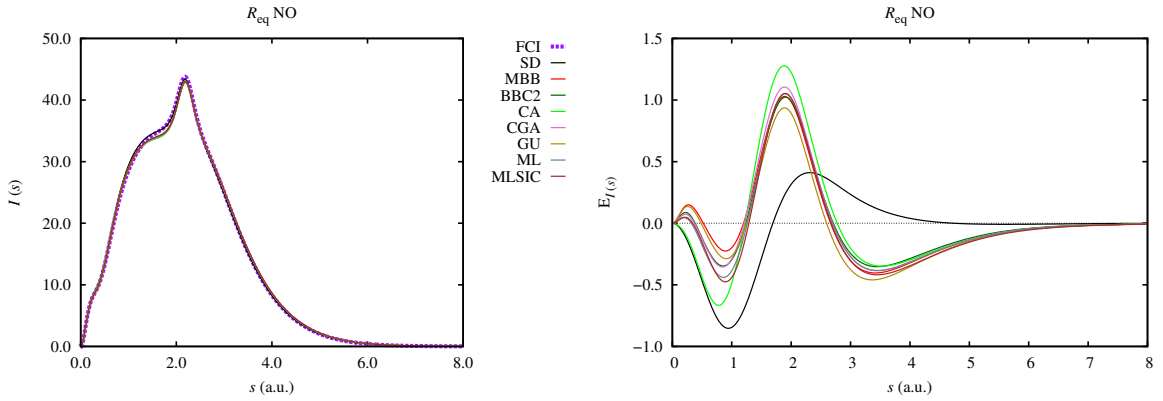


Figure S17: NO intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

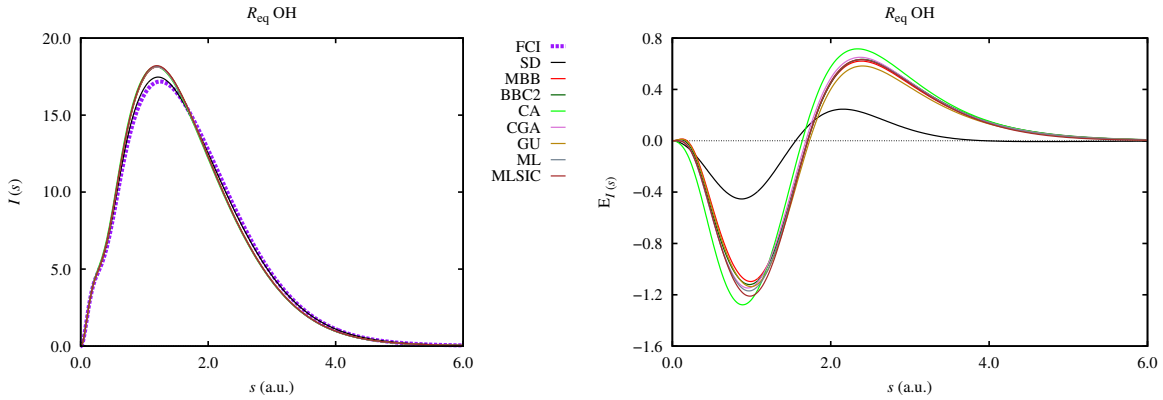


Figure S18: OH intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

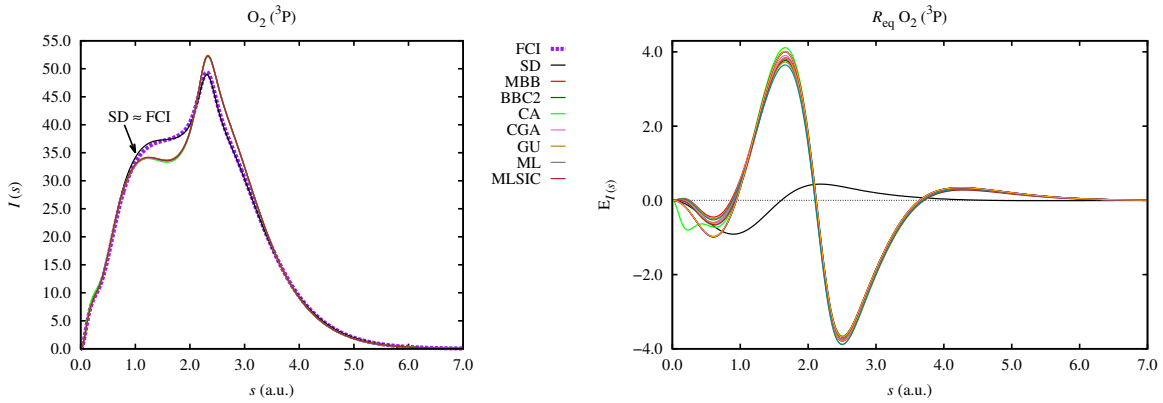


Figure S19: Triplet O_2 intracule probability densities (IPD) at the equilibrium geometry R_{eq}. See caption of Figure S1 for more details.

1.2 Stretched geometries $5R_{eq}$

1.2.1 Closed-shell molecules

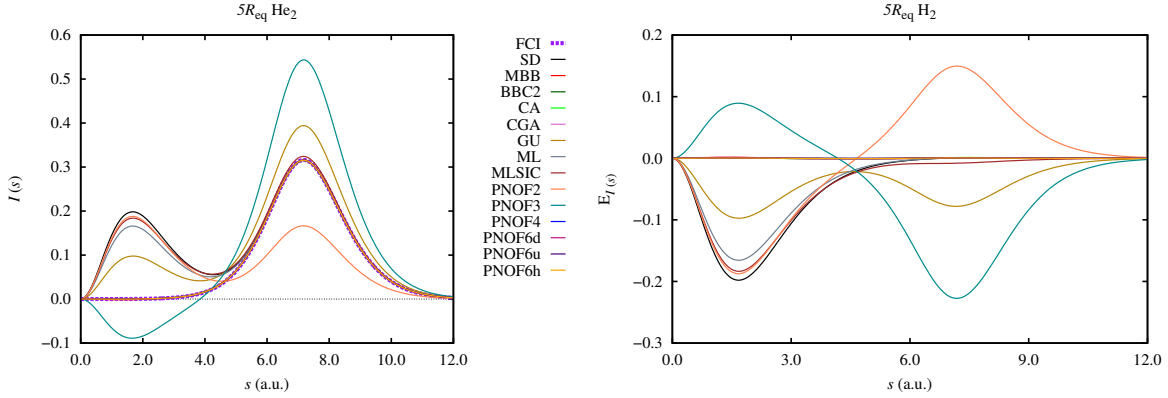


Figure S20: H_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

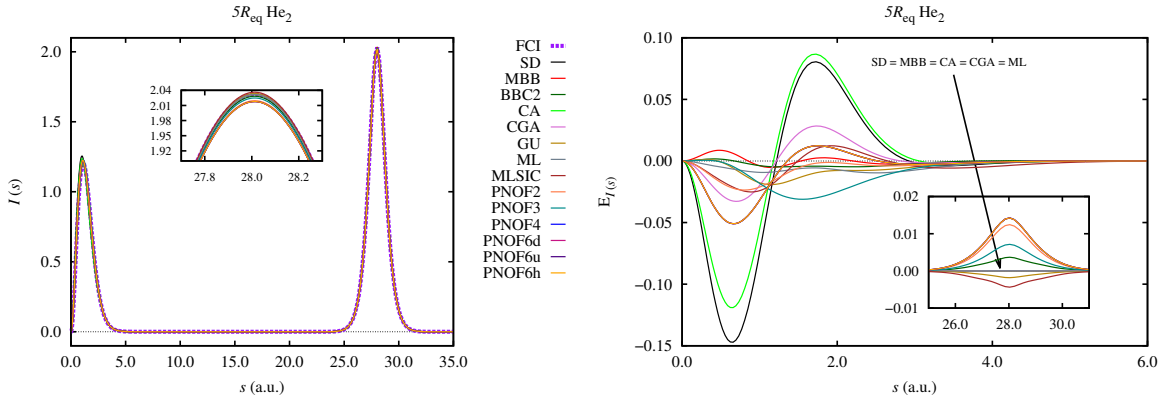


Figure S21: He_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

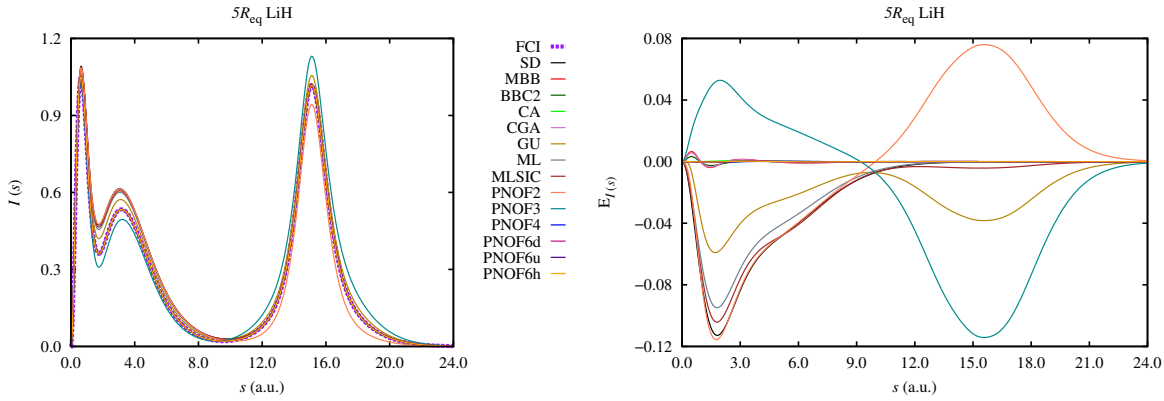


Figure S22: LiH intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

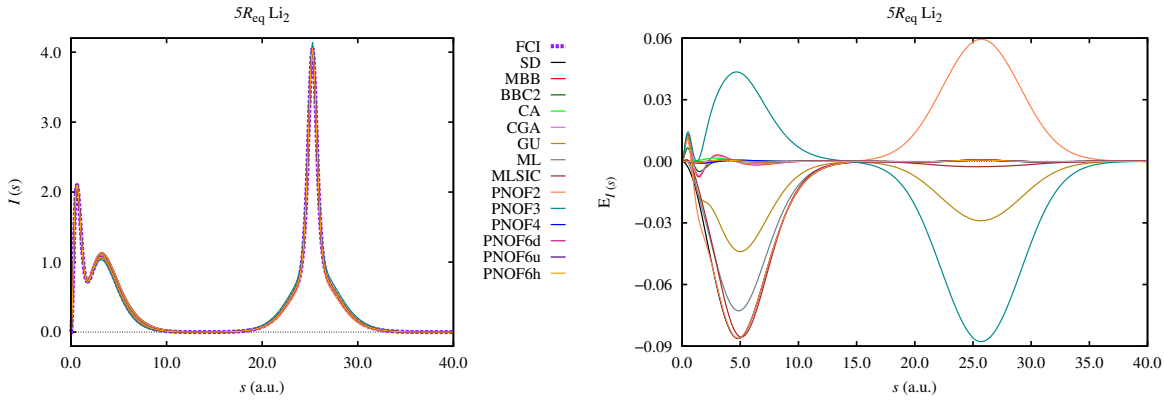


Figure S23: Li_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

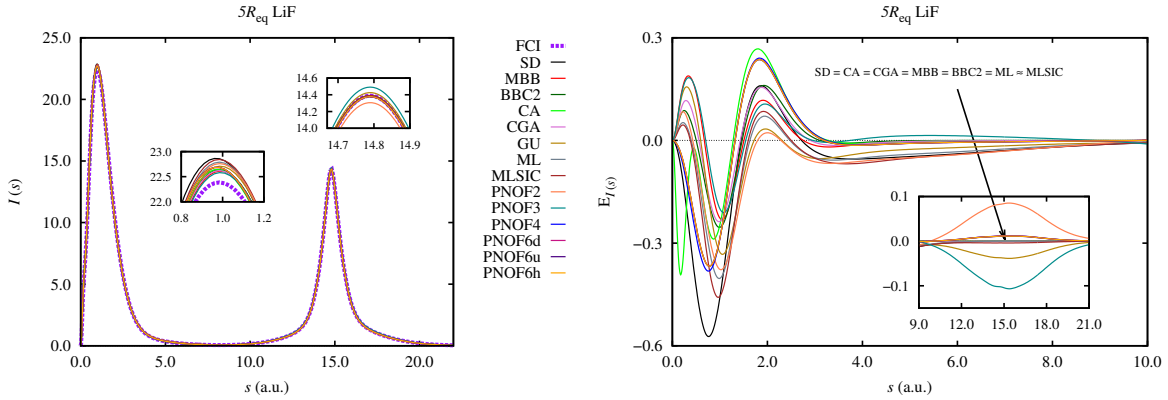


Figure S24: LiF intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

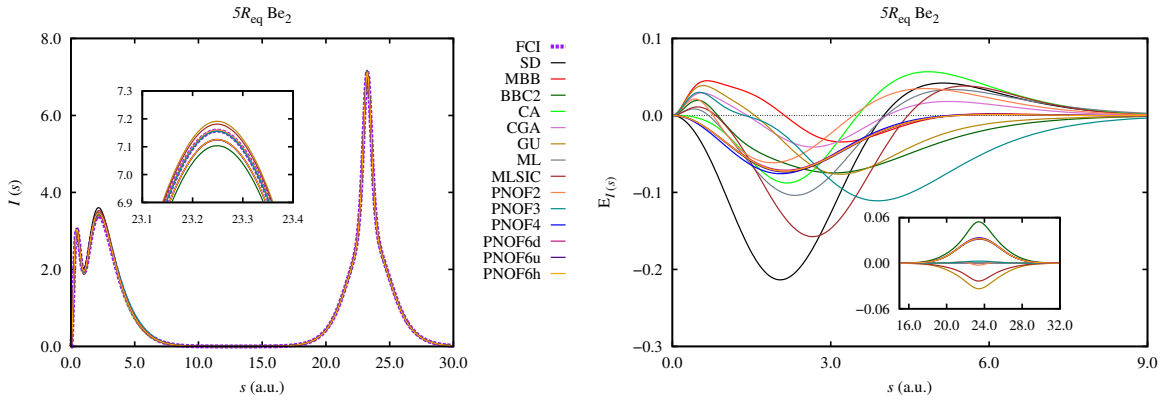


Figure S25: Be_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

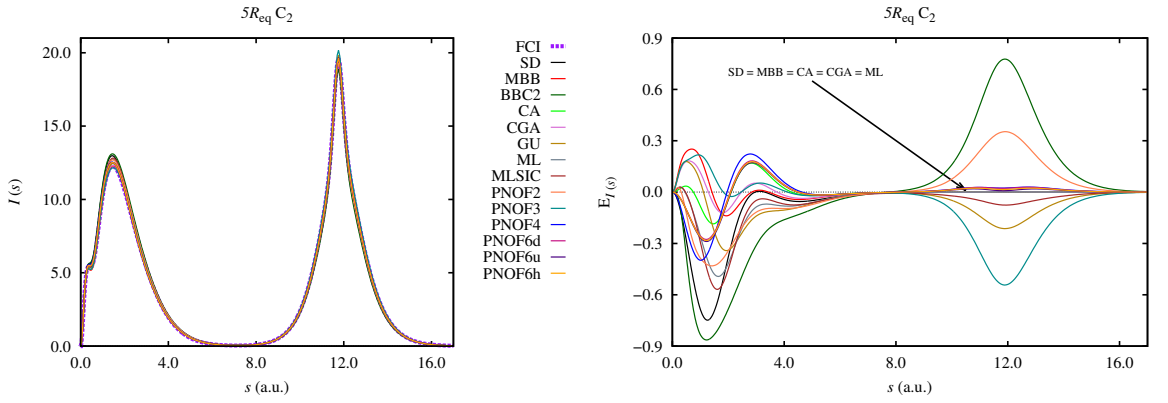


Figure S26: C_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

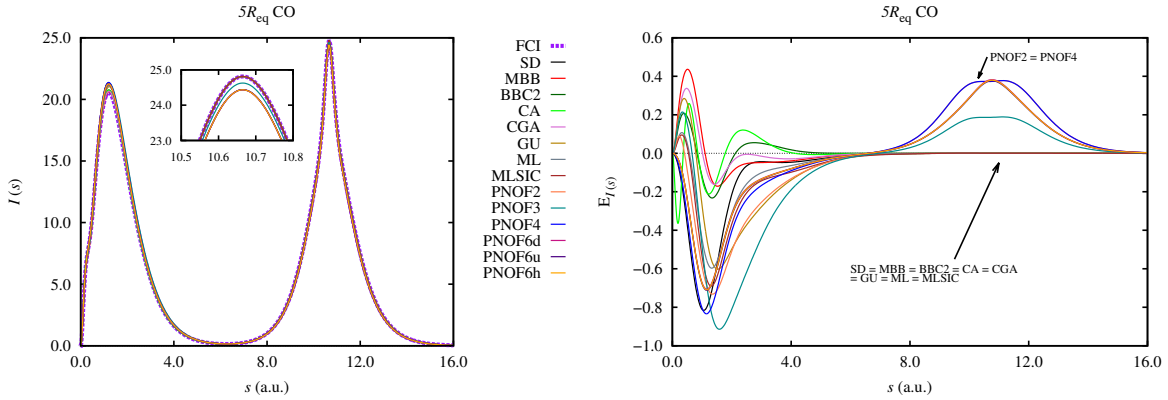


Figure S27: CO intracule probability densities (IPD) at $5R_{\mathrm{eq}}$. See caption of Figure S1 for more details.

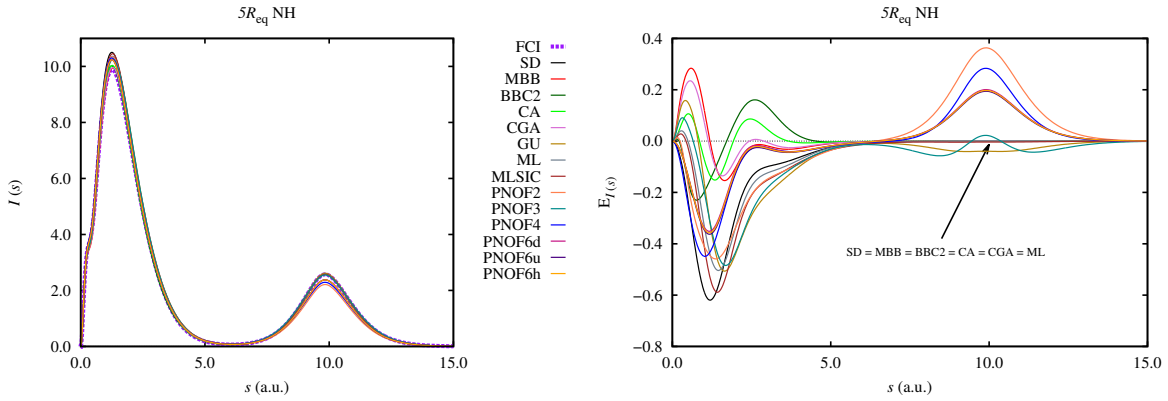


Figure S28: NH intracule probability densities (IPD) at $5R_{\mathrm{eq}}$. See caption of Figure S1 for more details.

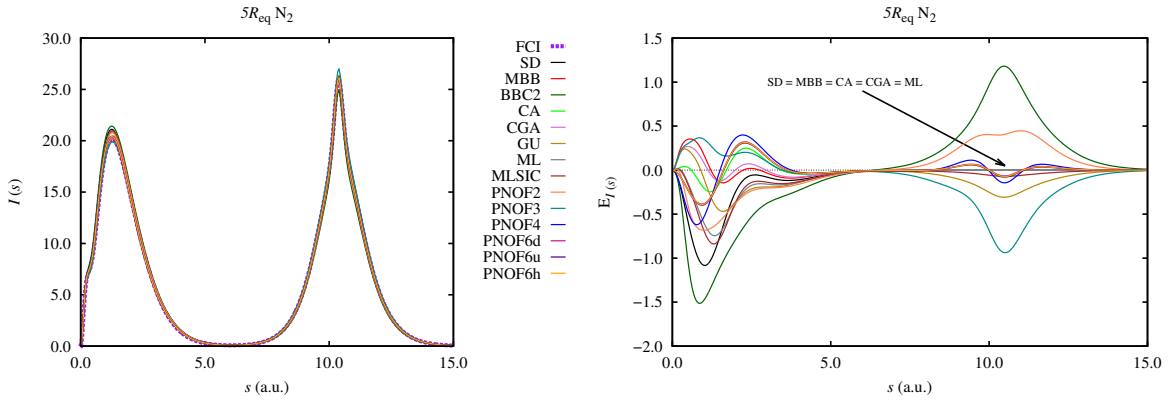


Figure S29: N_2 intracule probability densities (IPD) at $5R_{\mathrm{eq}}$. See caption of Figure S1 for more details.

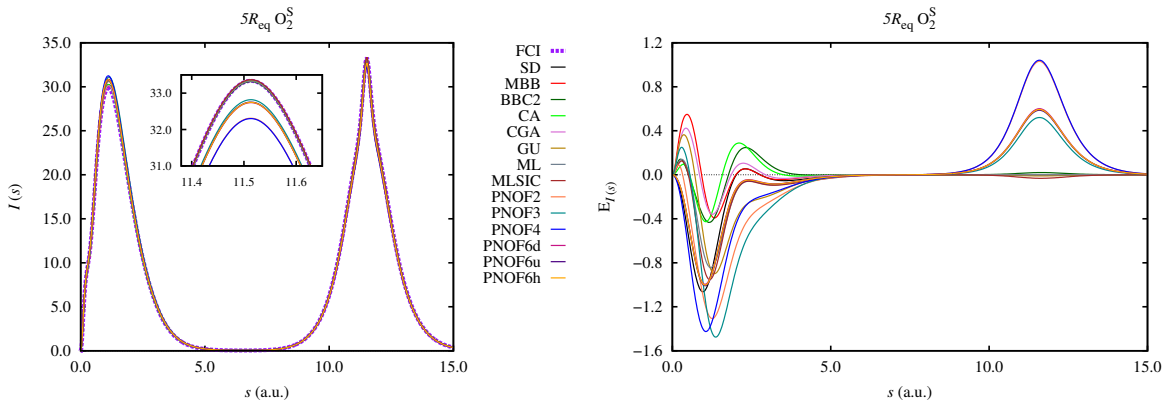


Figure S30: Singlet O_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

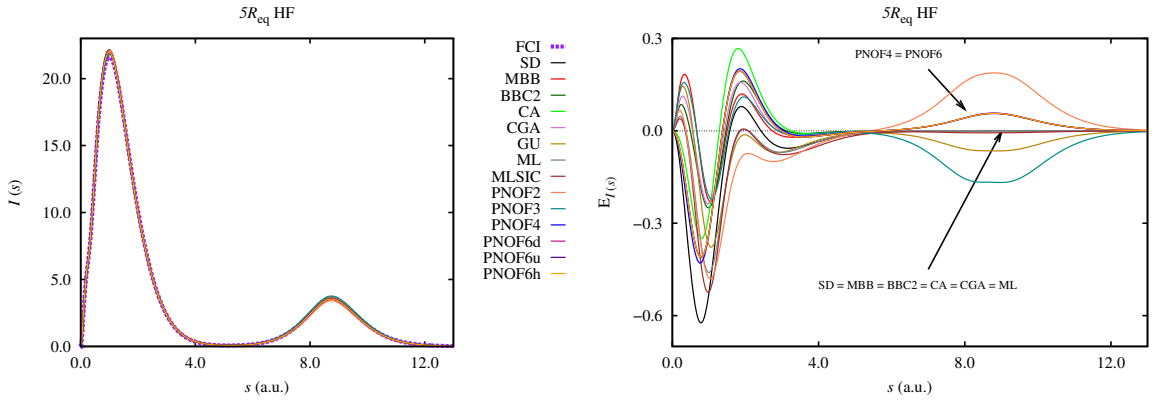


Figure S31: HF intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

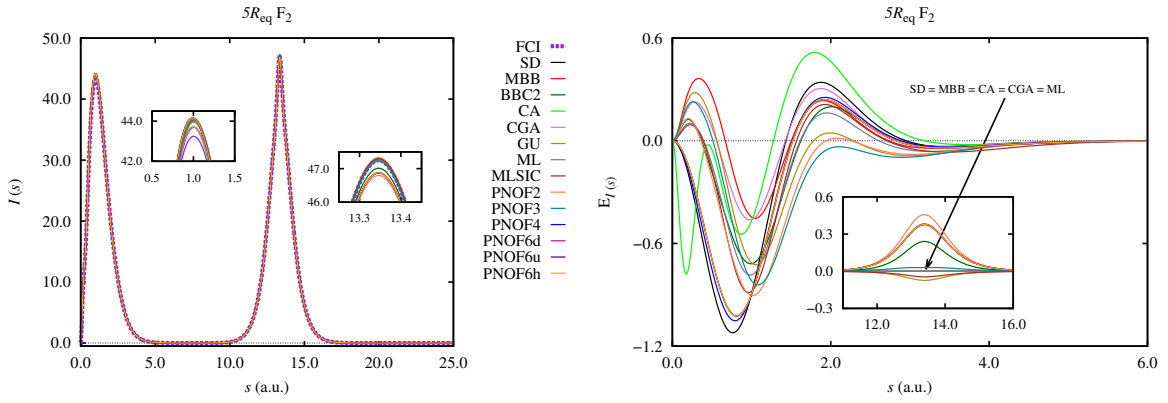


Figure S32: F_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

1.2.2 Open-shell molecules

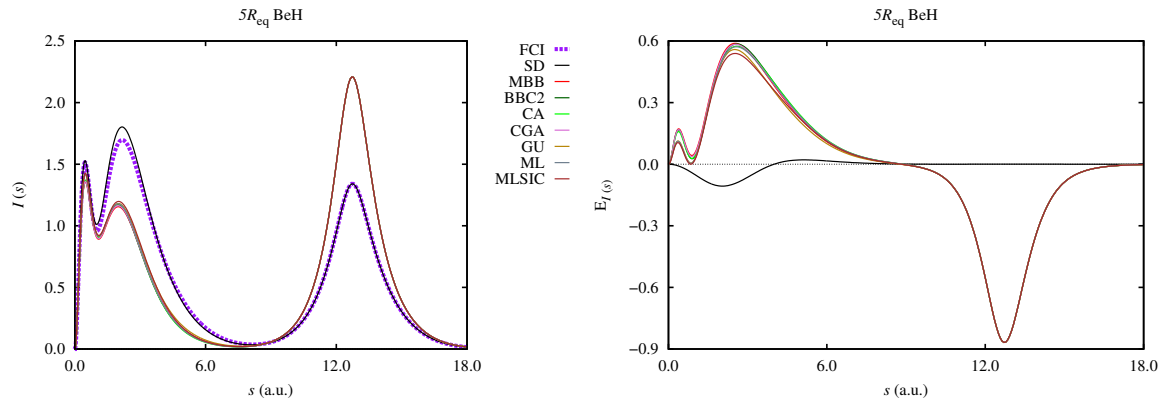


Figure S33: BeH intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

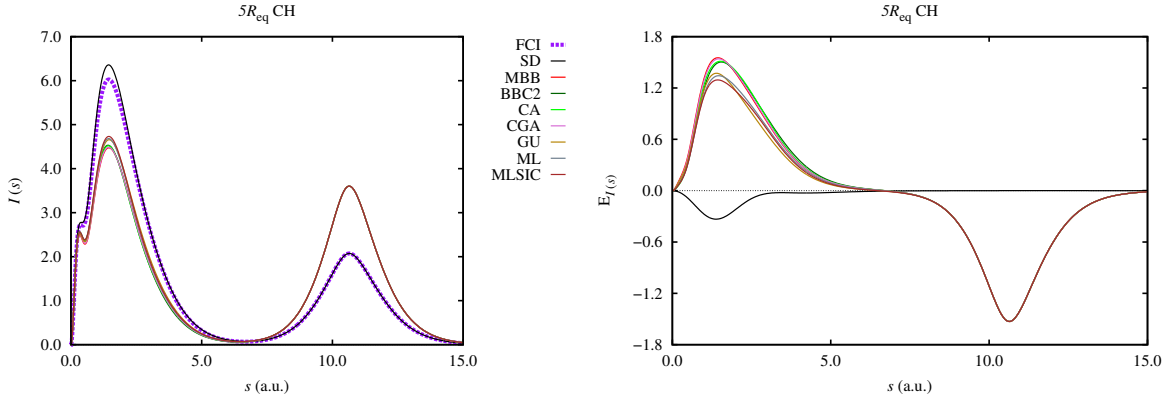


Figure S34: CH intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

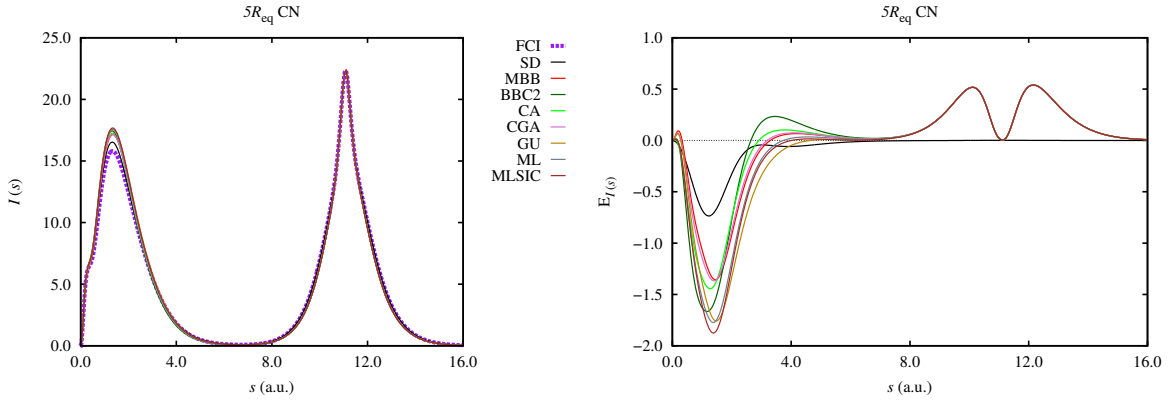


Figure S35: CN intracule probability densities (IPD) at $5R_{\mathrm{eq}}$. See caption of Figure S1 for more details.

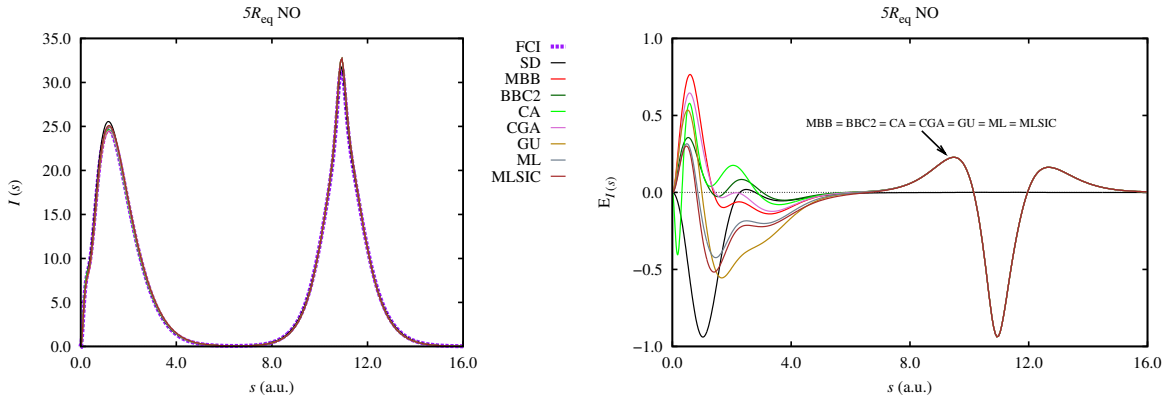


Figure S36: NO intracule probability densities (IPD) at $5R_{\mathrm{eq}}$. See caption of Figure S1 for more details.

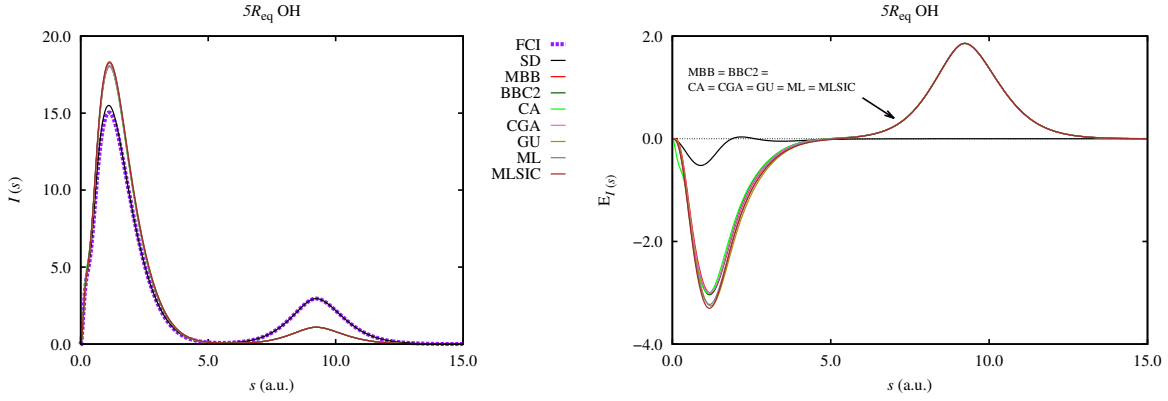


Figure S37: OH intracule probability densities (IPD) at $5R_{\mathrm{eq}}$. See caption of Figure S1 for more details.

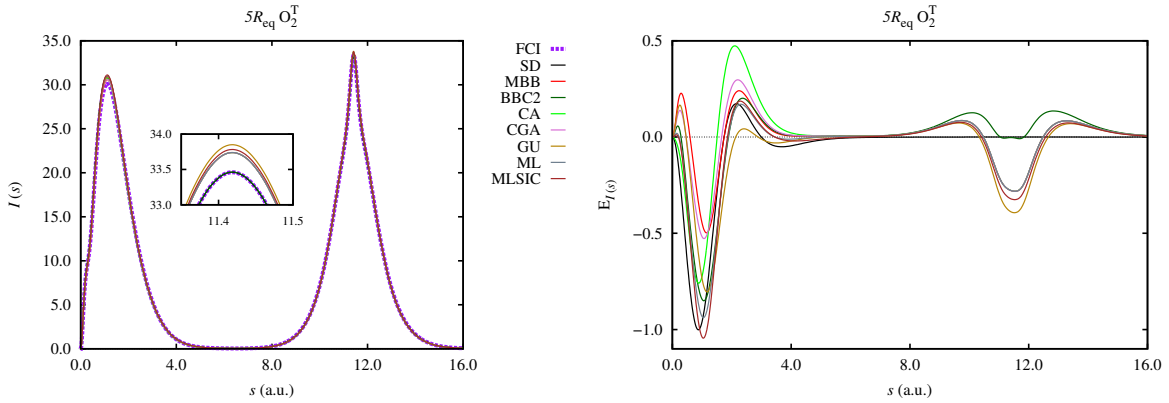


Figure S38: Triplet O_2 intracule probability densities (IPD) at $5R_{eq}$. See caption of Figure S1 for more details.

Chapter 6

Results and discussion

In this chapter, the results introduced in the works compiled in Chapters 3–5 are discussed and summarized. In Section 6.1, the ability to describe nondynamic and dynamic correlation effects of the correlation components introduced in Chapter 3 is validated. By means of textbook examples of dynamic and nondynamic correlation, as well as physical models and other molecules, the validation of the c_I and c_{II} correlation components is done in Sections 6.1.1 and 6.1.2, respectively. Insight is given into London dispersion interactions in Section 6.1.2.1. Particular profiles are detected in both correlation components of the Coulomb hole, according to the range considered. This is discussed in Section 6.1.3. An analysis of the spin components of the Coulomb hole and its partitions is analyzed in 6.1.4.

The new universal footprint found in the intracule pair density for dispersion interactions introduced in Chapter 4 is reviewed and expanded in Section 6.2. Finally, Section 6.3 encloses the outcome of the rDMFT benchmark introduced in Chapter 5.

6.1 On the pair density-based decomposition of the Coulomb hole

Chapter 3.1 presented the decomposition of the correlation part of the 2-PD, $\Delta\rho_2^{\mathrm{c}}$, into two components, named $\Delta\rho_2^{\mathrm{c}_I}$ and $\Delta\rho_2^{\mathrm{c}_{II}}$, where

$$\rho_2(\boldsymbol{r_1},\boldsymbol{r_2}) = \rho_2^{\mathrm{HF}}(\boldsymbol{r_1},\boldsymbol{r_2}) + \Delta\rho_2^{\mathrm{c}}(\boldsymbol{r_1},\boldsymbol{r_2}) = \rho_2^{\mathrm{HF}}(\boldsymbol{r_1},\boldsymbol{r_2}) + \Delta\rho_2^{\mathrm{c}_I}(\boldsymbol{r_1},\boldsymbol{r_2}) + \Delta\rho_2^{\mathrm{c}_{II}}(\boldsymbol{r_1},\boldsymbol{r_2}). \tag{6.1}$$

The definitions for the components arising from the splitting of $\Delta\rho_2^{\mathrm{c}}(\boldsymbol{r_1},\boldsymbol{r_2})$ are:

$$\Delta\rho_2^{\mathrm{c}_I}(\boldsymbol{r_1},\boldsymbol{r_2}) := \rho_2^{\mathrm{SD}}(\rho_1,\boldsymbol{r_1},\boldsymbol{r_2}) - \rho_2^{\mathrm{SD}}(\rho_1^{\mathrm{HF}},\boldsymbol{r_1},\boldsymbol{r_2}) \tag{6.2}$$

and

$$\Delta\rho_2^{c_{II}}(\boldsymbol{r_1},\boldsymbol{r_2}) := \rho_2(\boldsymbol{r_1},\boldsymbol{r_2}) - \rho_2^{\mathrm{SD}}(\rho_1,\boldsymbol{r_1},\boldsymbol{r_2}), \tag{6.3}$$

where $\rho_2^{\mathrm{SD}}(\rho_1^{\mathrm{HF}},\boldsymbol{r_1},\boldsymbol{r_2}) \equiv \rho_2^{\mathrm{HF}}(\boldsymbol{r_1},\boldsymbol{r_2})$ is the HF 2-PD (Eq. 1.37), $\rho_2^{\mathrm{SD}}(\rho_1,\boldsymbol{r_1},\boldsymbol{r_2})$ is the SD approximation of the 2-PD (Eq. 1.42), and $\rho_2(\boldsymbol{r_1},\boldsymbol{r_2})$ is the FCI 2-PD. Note that the definitions of $\Delta\rho_2^{c_I}$ and $\Delta\rho_2^{c_{II}}$ are analogous to the ones for Λ_{ij}^{kl} and Γ_{ij}^{kl} (Eqs. 1.86 and 1.87), respectively, presented in Section 1.4.2. The main difference is that Eqs. 6.2 and 6.3 are not restricted to the NOrb representation, and that the decomposition of $\Delta\rho_2^{\mathrm{c}}$ is exact.

The criterion to split $\Delta\rho_2^{\mathrm{c}}$ in such a way is based on the inability of the SD approximation of the 2-PD (Eq. 1.42) to describe dynamic correlation effects [136,222]. It is accepted that the SD 2-PD describes a large part of nondynamic correlation, but it is too bold to assume that it considers *all* of them. Because of this, we can assure that $\Delta\rho_2^{c_I}$ *only* contains nondynamic correlation effects, whereas $\Delta\rho_2^{c_{II}}$ mainly describes dynamic correlation, but also some nondynamic correlation effects. The hypothesis for this decomposition is analogous to the argument presented in Section 1.4.2: the 1-RDM (ρ_1) is able to capture the nondynamic correlation effects. Nondynamic correlation arises when going from a single configurational picture (that is, a CI vector mainly described by the HF determinant and a large number of lowly contributing configurations) to a multiconfigurational one (a CI vector composed by more than one dominant determinant besides the HF one). Because of this, going from one picture to another the 1-RDM will undergo large changes. Nondynamic correlation causes large and global changes in the electron density caused by the large mixing of configurations. Therefore, $\Delta\rho_2^{c_I}$ is close to zero in a system mainly described by the HF determinant (a system that mainly suffers of dynamic correlation effects). Instead, a strong mixing of configurations with the HF determinant (which gives rise to nondynamic correlation effects) implies a large value in $\Delta\rho_2^{c_I}$ because the 1-RDM is substantially different from the HF 1-RDM.

One of the main goals of this thesis is to study the range-separation of electron correlation and its components. Because the IPD (Eq. 1.71) summarizes the electron correlation effects contained in the 2-PD, it permits a natural separation of electron correlation by range. As well, the IPD is the building block to construct Coulomb holes (see Chapter 1.3.1); one can construct a correlation hole for each 2-PD correlation component. If the Coulomb hole is defined as the difference between the FCI (correlated reference) and the HF (uncorrelated reference) IPDs (Eq. 1.74),

$$h_{\mathrm{c}}(s) = I(\Delta\rho_2^{\mathrm{c}},s) = I(\rho_2^{\mathrm{FCI}},s) - I(\rho_2^{\mathrm{HF}},s), \tag{6.4}$$

it is sensible to apply the correlation splitting for the 2-PD into it:

$$h_{\mathrm{c}}(s) \equiv I(\Delta\rho_2^{\mathrm{c}}, s) = I(\Delta\rho_2^{\mathrm{c}_I}, s) + I(\Delta\rho_2^{\mathrm{c}_{II}}, s) \equiv h_{\mathrm{c}_I}(s) + h_{\mathrm{c}_{II}}(s). \tag{6.5}$$

As we shall see, Eq. 6.5 permits the study, identification and treatment of each correlation type separately and by range.

6.1.1 h_{c_I} and nondynamic correlation

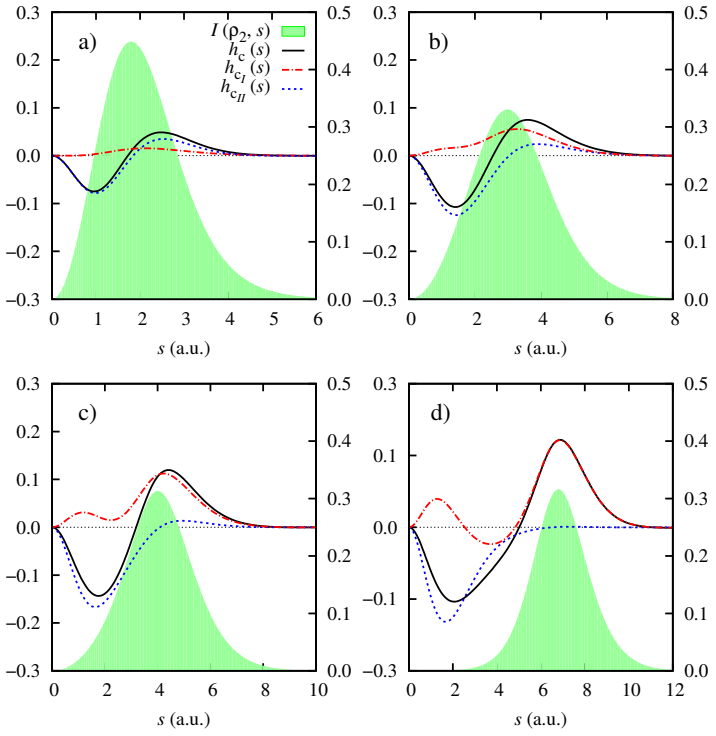


Figure 6.1: Coulomb hole and its correlation components for H_2 at a) 1.32, b) 2.83, c) 3.78, and d) 6.61 a.u. bond distances. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison.

The basis of the decomposition relies on the ability of the 1-RDM to capture nondynamic correlation effects, but not retrieving dynamic correlation. Although the definition of the 2-PD c_I component (Eq. 6.2) is a difference between two 2-PDs, both ρ_2^{SD} and ρ_2^{HF} use the same expression (same "recipe") but a different input variable, the 1-rDM (two different "ingredients"). What is actually evaluated in Eq. 6.2 is, then, the dissimilarity between the FCI and HF 1-rDMs by means of the 2-PD. Expressed in terms of NOrbs and NOccs, it evaluates the difference between the FCI occupancies (fractional numbers compressed between 1 and 0) and the HF ones (1 or 0). When a molecule requires a multiconfigurational wavefunction, nondynamic correlation arises between electron pairs. Cases in which the HF wavefunction fails

to describe an electronic system are the stretching of a bond and molecular dissociations; or, in general, orbital degeneracies caused by the symmetry of the system. Therefore, molecular dissociations and cases with orbital degeneracies in the ground state are considered in this section to evaluate whether $\Delta\rho_2^{c_I}$ captures nondynamic correlation effects.

6.1.1.1 Chemical examples

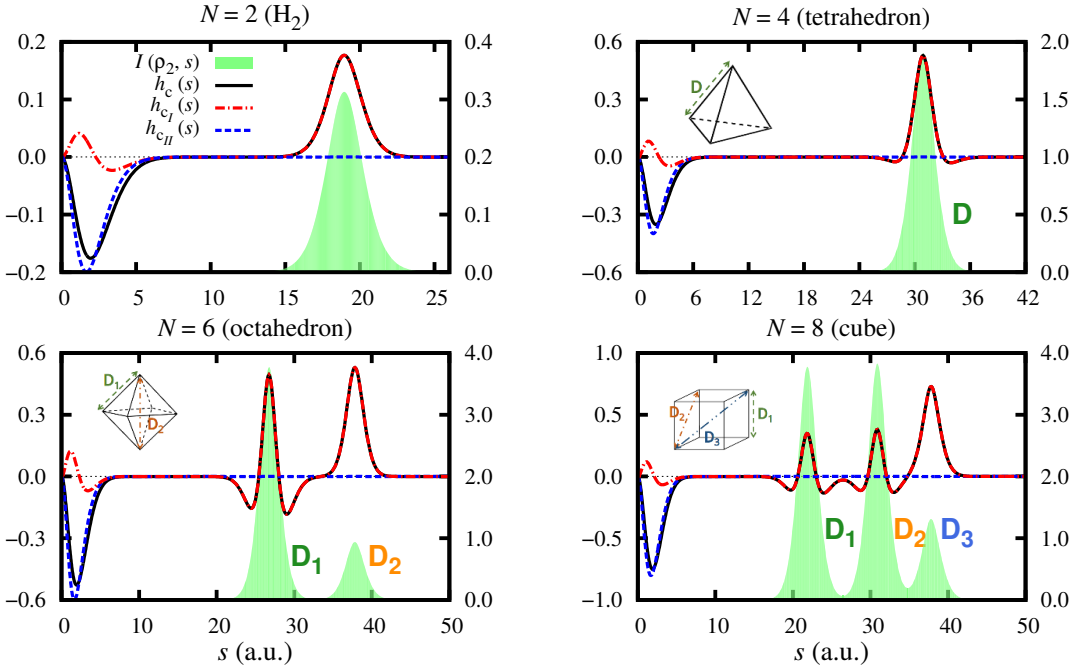


Figure 6.2: Coulomb hole h_c in black, and its c_I (red) and c_{II} (blue) correlation components of H_N with $N = 2, 4, 6$ and 8. The FCI IPD, $I(\rho_2^{\text{FCI}}, s)$ is represented in green (right y-axis) for comparison.

The stretching of the H–H bond in H_2 is the most elemental example to evaluate $\Delta\rho_2^{c_I}$. It is well known that a single Slater determinant is not able to describe the homolytic separation of two electrons into their respective hydrogen nuclei. Fig. 6.1 contains the Coulomb holes and its components at different bond distances. It can be seen that, as the bond stretches, $h_{c_{II}}$ becomes less important and h_{c_I} turns into the leading component of the Coulomb hole. Another example is polyhedral H_N, a system composed of $N = 2, 4, 6, 8$ hydrogen atoms equidistantly separated a distance of 18.9 a.u. from the center of mass, forming a polyhedron. Fig. 6.2 shows that Coulomb holes of H_N present a large c_I hole that peaks around the interfragment distances with the other hydrogens (labeled with a D in the figure).

Because nondynamic correlation rises when a molecule dissociates, h_{c_I} predominates in the LR region, caused by the LR interactions between the electrons found in each atomic center. Through the molecules analyzed, h_{c_I} is usually positive at both SR and LR regions, being mostly long-ranged with large values due to the small electron pair probability of the HF 2-PD at those ranges. Exceptions to this statement are found in molecules where HF is not able to produce the correct homolytic dissociation, such as in LiH and H_4 at the C_{3v} symmetry point group, presented in Chapter 3.2. See Fig. 6.3 for the LiH Coulomb hole as an example. In these molecules, HF dissociates into fragments with a noninteger number of electrons. This causes some fragments to have a larger electron pair population, and, therefore, h_{c_I} is negative in the LR region. Instead, if HF would dissociate into the correct number of electron pairs in each fragment (as in the UHF method), h_{c_I} would be positive in the LR region of the hole (see Figs. S5 and S6 in Chapter 3.2).

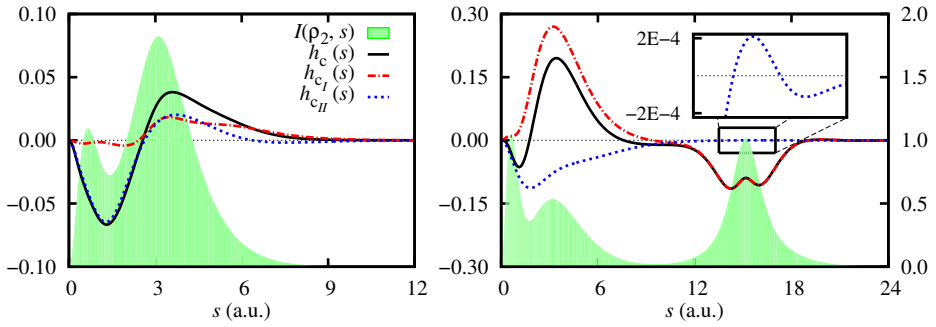


Figure 6.3: Coulomb hole h_c in black, and its c_I (red) and c_{II} (blue) correlation components of LiH at the equilibrium geometry at 3.02 a.u., and 15.12 a.u. bond length. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison. The inset plot shows the effect of dispersion interactions in the LR region of $h_{c_{II}}$.

On the other hand, note that h_c, h_{c_I}, and $h_{c_{II}}$ have nonzero values in the SR region at the stretched geometries of the H_N molecules. Because there exist no SR interactions within a hydrogen atom (since there is only one electron; see the shape of the FCI IPD in the green shadowed region in Figs. 6.1, 6.2, for instance), these nonzero values in the SR region of the holes are consequence of the SR description of the HF 2-PD. This amount corresponds to the number of electron pairs that the HF description places at short interelectronic distances.

H_4 in D_{2h} and D_{4h} symmetry point group is a convenient system to analyze for it permits tuning the effects of electron correlation in terms of the distance between the center of mass and a hydrogen atom R, and the angle θ formed between the center of mass and two adjacent hydrogens. Cases where R is large share the same profiles

with the already discussed H_2 and H_N molecules: h_{c_I} is large and positive in the LR region of the hole, centered around the interfragment distance with other hydrogens, and caused by the nondynamic correlation rising from molecular dissociation. In the particular case of H_4 in the D_{4h} symmetry point group, an important orbital degeneracy exists caused by the symmetry of the molecule. This orbital degeneracy causes SR interactions between electron pairs that are caused by nondynamic correlation effects. The c_I hole component is positive in the SR region of the Coulomb hole, indicating that $\Delta\rho_2^{c_I}$ captures some of the SR nondynamic correlation effects from $\Delta\rho_2^{c}$.

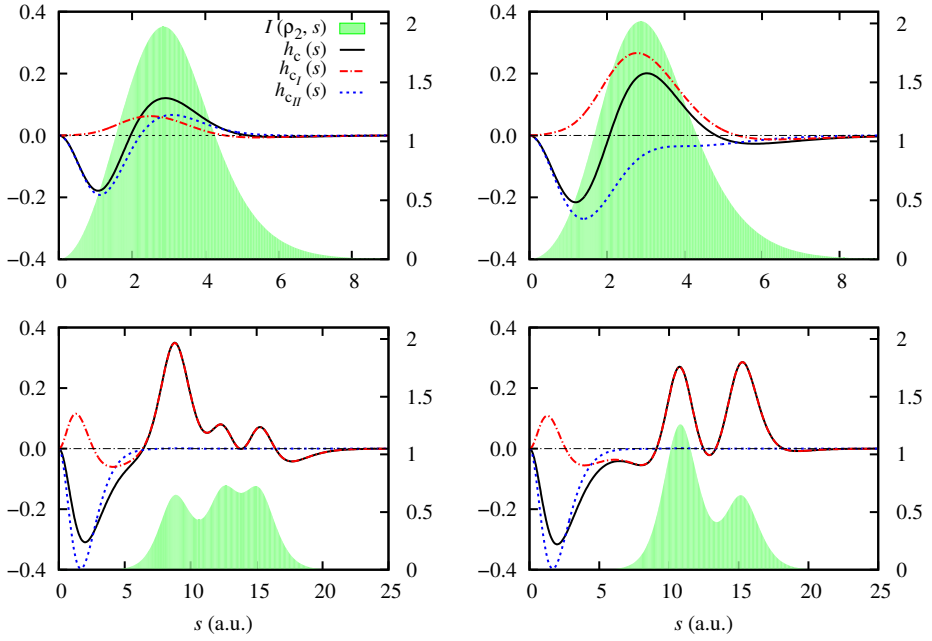


Figure 6.4: Coulomb hole and its correlation components for four conformations of D_{2h}/D_{4h} H_4. a) $R = 1.51$ a.u. and $\theta = 0.39\pi$ (D_{2h}), b) $R = 1.51$ a.u. and $\theta = \pi/2$ (D_{4h}), c) $R = 7.59$ a.u. and $\theta = 0.39\pi$ (D_{2h}), and d) $R = 7.59$ a.u. and $\theta = \pi/2$ (D_{4h}). R is the distance from the center of mass to one hydrogen, and θ is the angle between the center of mass and the two closest hydrogen atoms. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison.

The beryllium atom is a textbook case of nondynamic correlation due to the near-degeneracy of the $2s$ and $2p$ orbitals, a feature that is passed to the beryllium dimer (for more details of Be atom, see Fig. 6.9 for the Coulomb hole, and Section 6.1.1.3 for an analysis of the isoelectronic series of Be). $h_{c_{II}}$ is not zero when Be_2 is at the equilibrium geometry, due to the SR nature of the interactions that give rise to dynamic correlation. However, h_{c_I} is larger and is the leading component that defines the Coulomb hole (see Fig. 6.5). Because the SR interactions between electron pairs delocalized in the $2s - 2p$ orbitals give rise to nondynamic correlation, $\Delta\rho_2^{c_I}$ describes such interactions and are reflected in the magnitude of h_{c_I}. The SR region

of h_c in the dissociated geometry of Be_2 is equivalent to the picture obtained for the equilibrium case since this SR region describes the interactions within a beryllium atom. Certainly, the correlated motion of electrons in Be is nondynamic because of the $2s-2p$ orbital degeneracy, contemplated by $\Delta\rho_2^{c_I}$ and reflected in h_{c_I}. The maximum in the LR region of the Coulomb hole is characterized by h_{c_I}, caused by the orbital degeneracy arising from the bond stretching of Be_2. As in H_2, the h_{c_I} peak is positive due to the covalent nature of the bond. Despite $h_{c_{II}}$ seems to be zero at this region of the hole, it presents a small contribution caused by dispersion forces (see Section 6.1.2.1 and Fig. 6.12).

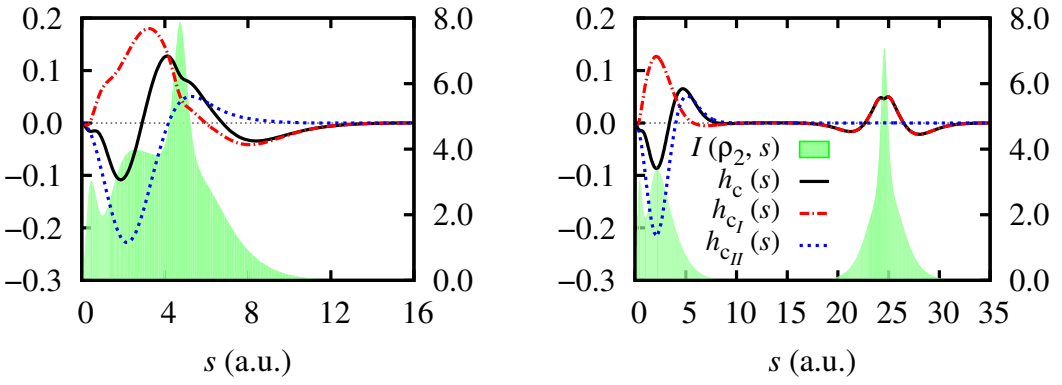


Figure 6.5: Coulomb hole and its correlation components for Be_2 at 4.72 a.u. (left) and 24.57 a.u. (right) bond distance. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison. Be_2 has been calculated with the FCI(FC)/aug-cc-pVTZ level of theory. See Fig. 6.12 for a zoomed detail on the LR region of the stretched geometry.

6.1.1.2 Physical models

The Hubbard model is widely used in physics, for it brings the ability to tune correlation that can be easily computed [223, 224]. Particularly, the Hubbard dimer contains two interacting electrons in two orbitals found in a lattice. The Hamiltonian of the Hubbard model tunes the strength of the interelectronic interactions with the on-site interaction U and the hopping t parameters:

$$\hat{H} = -t \sum_{\langle\mu,\nu\rangle,\sigma} (\hat{a}^\dagger_{\mu\sigma}\hat{a}_{\nu\sigma} + \hat{a}^\dagger_{\nu\sigma}\hat{a}_{\mu\sigma}) + U \sum_{\mu} \hat{\rho}_{\mu\alpha}\hat{\rho}_{\mu\beta}, \tag{6.6}$$

where μ and ν denote the two sites of the lattice, σ the spin of the electron, $\hat{a}^\dagger_{\mu\sigma}$ and $\hat{a}_{\mu\sigma}$ are the creation and annihilation operators of one electron with spin σ in site μ, and $\hat{\rho}_{\mu\sigma}$ is the one-particle number operator with spin σ on site μ. Large U/t values induce nondynamic correlation to the system, and small U/t values result in a dimer

with only dynamic correlation. Fig. 6.6 contains four different U/t values of the real-space Hubbard dimer. When U/t is small, an electron pair can be located in the same site of the lattice. This scenario is described by $h_{c_{II}}$. However, an increase in U/t causes a decrease in $h_{c_{II}}$ and the Coulomb hole is taken over by h_{c_I}, due to the strong repulsion felt by the electron pair, now found at different sites. This example also features the ability of $\Delta\rho_2^{c_I}$ to describe nondynamic correlation effects.

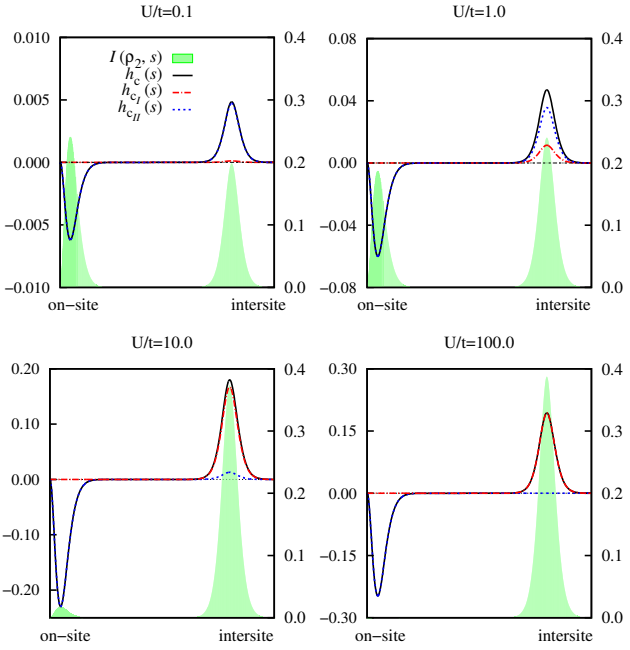


Figure 6.6: Coulomb hole and its correlation components for the half-filled Hubbard dimer in real space, with different U/t correlation parameters. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison.

The strength of electron correlation can also be modulated through the harmonium atom [225–227]. This model system is described with a Hamiltonian that approximates the interelectron repulsion with a harmonic potential:

$$\hat{H} = -\frac{1}{2}(\nabla_1^2 + \nabla_2^2) + \frac{1}{2}\omega^2(r_1^2 + r_2^2) + \frac{1}{|\boldsymbol{r_2} - \boldsymbol{r_1}|} \tag{6.7}$$

The strength of the repulsion is tuned with the confinement strength parameter ω: large ω values correspond to very small correlation, and strongly correlated cases are reached for very small ω values. The two-electron harmonium permits a study of varying electron correlation regimes with very small computational effort. Fig. 6.7 depicts the harmonium hole in three correlation regimes. At $\omega = 1000$ (the less correlated case), electrons do not suffer a large repulsion, so the dominant correlation type in the system is dynamic. This is the behavior described by $h_{c_{II}}$, that

is completely identical to the shape of the Coulomb hole, and h_{c_I} is flat along the x-axis. At $\omega = 1.0$ (middle graph in Fig. 6.7), h_{c_I} is not zero anymore. However, its contribution is considerably small compared to the one of $h_{c_{II}}$. h_{c_I} prevails over the c_{II} hole in the strongly correlated case of $\omega = 0.3$. Note that the average interelectronic distance $\langle s \rangle$ is larger in this case than in the other two considered (compare the x-axes in Fig. 6.7). Repulsion is so strong in this case that electron pairs are largely separated, giving rise to LR nondynamic correlation. These interactions are described by $\Delta\rho_2^{c_I}$.

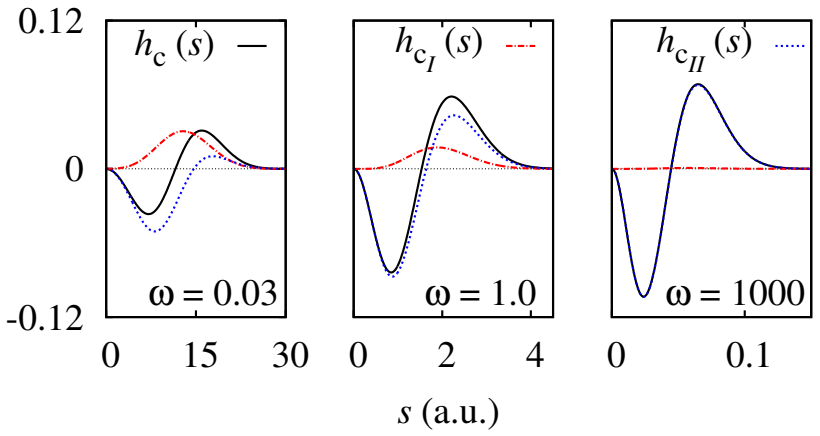


Figure 6.7: Coulomb hole and its correlation components for the two-electron harmonium atom, with different values of confinement strength: (left) $\omega = 0.3$ (center) $\omega = 1.0$ (right) $\omega = 1000$.

6.1.1.3 Types A and B nondynamic correlation

The isoelectronic series of beryllium, Be(Z), is an example in which nondynamic correlation dominates the electron correlation effects. The small $2s - 2p$ orbital energy gap in Be gives rise to nondynamic correlation. Hollett and Gill indicate in their work that Be(Z) is a series described by type B nondynamic correlation and, because of that, the unrestricted formalism of HF is not able to describe it [35] (see Section 1.1.1 for more details about type A and type B nondynamic correlation). The ability of our correlation decomposition to discern between both types of nondynamic correlation is considered in this Section.

Type A nondynamic correlation arises in molecules with absolute near-degeneracies. The bond cleavage of H_2 (or the dissociation of a covalent bond in general) gives rise to type A. As it has been discussed through this section, $\Delta\rho_2^{c_I}$ describes the nondynamic correlation interactions arising from molecular dissociations, so type A nondynamic correlation should be contemplated in this electron correlation compo-

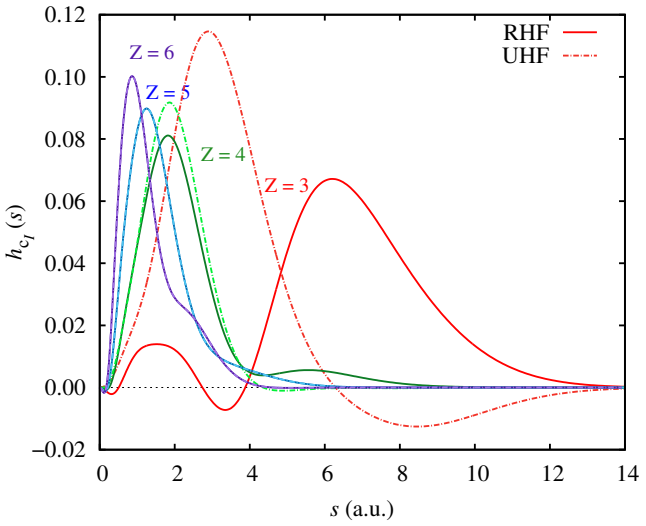


Figure 6.8: The c_I hole component of the isoelectronic series of beryllium, Be(Z). h_{c_I} for Z = 7 and 8 are redundant and therefore not represented. Solid (dashed) lines indicate h_{c_I} using the RHF (UHF) IPD as the uncorrelated reference, being $h^{\mathrm{R}}_{c_I}(s) = I(\rho_2^{\mathrm{SD}}, s) - I(\rho_2^{\mathrm{RHF}}, s)$ and $h^{\mathrm{U}}_{c_I}(s) = I(\rho_2^{\mathrm{SD}}, s) - I(\rho_2^{\mathrm{UHF}}, s)$.

Z	$\mathrm{E_{RHF}}$	$\mathrm{E_{UHF}}$	ΔE	$\langle S^2 \rangle_{\mathrm{UHF}}$	$\Delta\epsilon_{\mathrm{gap}}$	$\Delta\,(\Delta\epsilon_{\mathrm{gap}})$
3	-7.380119	-7.390860	-0.01074	0.7132	0.21864	0.15874
4	-14.57022	-14.57059	-0.00036	0.1378	0.37738	0.22385
5	-24.23339	-24.23339	0.0	0.0	0.60123	0.22636
6	-36.39601	-36.39601	0.0	0.0	0.82759	0.22368
7	-57.07217	-57.07217	0.0	0.0	1.05127	0.21915
8	-68.22939	-68.22939	0.0	0.0	1.27042	-

Table 6.1: The RHF and UHF energies, their energy difference, the spin contamination from the UHF calculation, the HOMO and LUMO energy gap, and the difference between two consecutive energy gaps for Be(Z) species with $3 \leq Z \leq 8$. Adaptation from the data presented in Chapter 3.2.

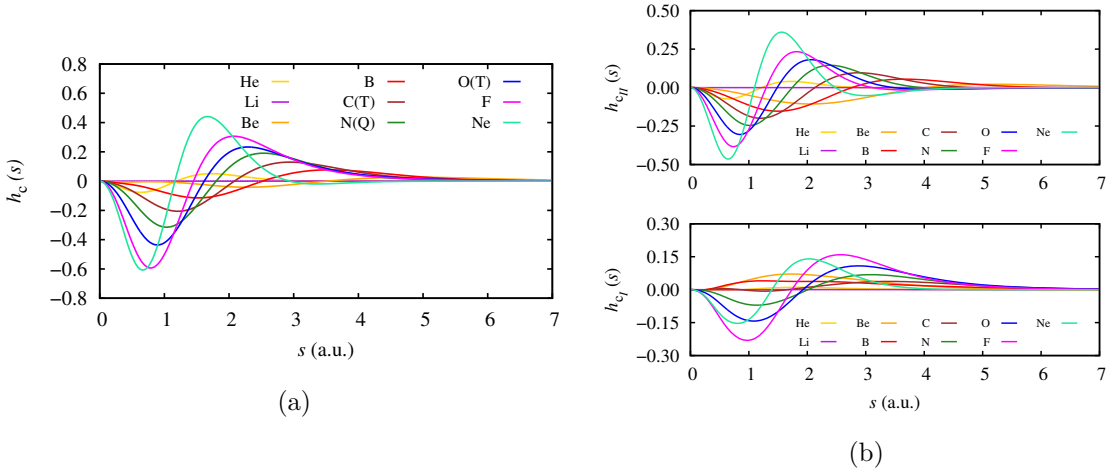
nent.

On the other hand, type B nondynamic correlation arises from relative near-degeneracies, and UHF is not able to describe its effects. Be(Z) is used to determine whether the correlation decomposition is able to discern between types A and B. The last column in Table 6.1 compiles the difference of HOMO-LUMO gaps for

two consecutive Be(Z), $\Delta(\Delta\epsilon_{\text{gap}})$. Because these differences are constant, Be(Z) presents a relative near-degeneracy with Z. Energies gathered in Table 6.1 indicate that the UHF method becomes less able to describe type B nondynamic correlation when the core potential Z increases with fixed N. UHF gives a different description from RHF at $Z = 3$ and $Z = 4$. Instead, from $Z = 5$ and beyond, the UHF description converges to the RHF case, indicating the inability of UHF to account for type B nondynamic correlation effects (a conclusion already presented in the work of Hollett and Gill [35]). h_{c_I} calculated using the UHF and RHF 2-PDs is different at $Z = 3$ and $Z = 4$ (see Fig. 6.8), and there is only one possible h_{c_I} for $Z = 5$ and beyond because the UHF wavefunction is the RHF one. With this, the inability of UHF to describe type B nondynamic correlation is also reflected in $\Delta\rho_2^{\text{c}_I}$.

6.1.2 $h_{\text{c}_{II}}$ and dynamic correlation

In contrast to $\Delta\rho_2^{\text{c}_I}$, the definition of $\Delta\rho_2^{\text{c}_{II}}$ does not rely on the 1-rDM differences but on the validity of the SD 2-PD to approximate the 2-PD (Eq. 6.3). The definition proposed for $\Delta\rho_2^{\text{c}_{II}}$ is actually coincident with the one for the cumulant of the density matrix (Eq. 1.32), which is commonly understood as the part of the 2-DM that contains the correlation effects in the 2-rDM [124, 125, 138, 147, 160]. Instead, $\Delta\rho_2^{\text{c}_{II}}$ is expected to contain the dynamic correlation effects and the remaining nondynamic correlation that the SD 2-PD is not able to retrieve (the amount that $\Delta\rho_2^{\text{c}_I}$ does not describe).



(a) (b)

Figure 6.9: (a) Coulomb holes of the He-Ne series in their ground state. (b) c_{II} (above) and c_I (below) components of the Coulomb hole. The sum of both c_I and c_{II} components recovers h_{c} shown in figure (a).

To corroborate the expected behavior of $\Delta\rho_2^{c_{II}}$, consider the atomic series compressed between helium and neon. The increase in the number of electrons implies a proportional growth of dynamic correlation. Coulomb holes of each atom are depicted in Fig. 6.9a, where, because of the increase of electron correlation with the number of electrons N within the same row, the magnitude of the holes increases. However, such growth does not have a specific trend and Coulomb holes present a particular shape according to the atom considered. Otherwise, $h_{c_{II}}$ shows a systematic growth with N within the same row, as it is exposed in Fig. 6.9b. In addition, a systematic shrinking of the hole is seen, and consequently a decrease of the average interelectronic distance $\langle s \rangle$, caused by the increasing attracting core potential. The systematic growth of $h_{c_{II}}$ is in agreement with the systematic increase of dynamic correlation felt for an atomic series within the same row of the periodic table. Contrarily, no specific trend is observed for h_{c_I}. Because nondynamic correlation is not universal and system-dependent, its manifestation depends on the nature of each atom. This system-dependent particularity of nondynamic correlation is represented correctly by h_{c_I} in the atomic series, and changes according to the importance of nondynamic correlation in these systems.

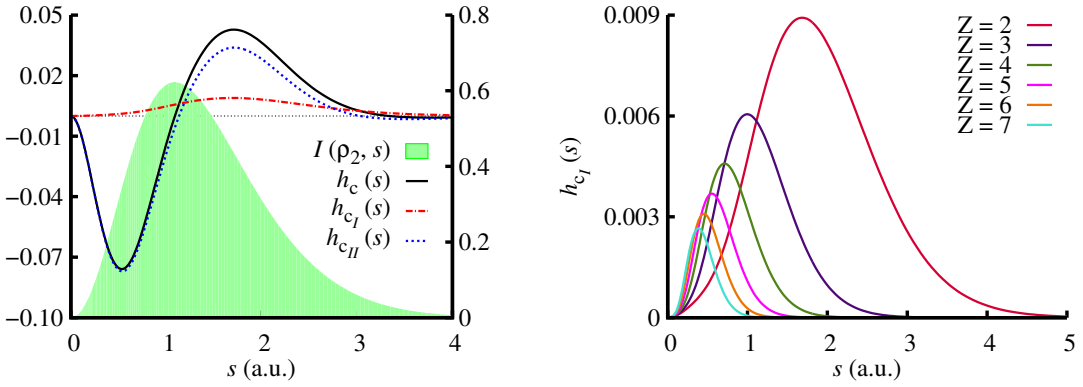


Figure 6.10: (Left) Coulomb hole and its correlation components for helium. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison. (Right) The c_I hole component for the isoelectronic series of helium comprehending $2 \leq Z \leq 7$.

The isoelectronic series of helium, He(Z), is a dynamic-correlation dominated series in which an increase in the effective potential Z when the number of electrons N is kept constant results into a shrinking of the electron density and, consequently, electron pairs undergo more important dynamic correlation effects. When the core potential is large compared to the number of electrons, the HF determinant describes the electron density rather correctly, and, consequently, the presence of nondynamic correlation in the system decreases. The h_{c_I} and $h_{c_{II}}$ profiles obtained for the series

are in agreement with the behavior described for dynamic and nondynamic correlation. Fig. 6.10 shows how h_{c_I} diminishes with increasing Z. Because $\Delta\rho_2^{c_I}$ becomes zero, the correlation decomposition indicates that He(Z) is a series mainly described by dynamic correlation.

Through all the systems analyzed in this thesis, it can be seen that $h_{c_{II}}$ mainly describes the SR region of the Coulomb hole, unless short-ranged interactions that give rise to nondynamic correlation occur (as in Be_2, for instance). The nature of dynamic correlation is mainly short-ranged, since electrons repel each other as a consequence of being charged and fermionic particles. Therefore, it is not surprising that $h_{c_{II}}$ is essentially short-ranged. Conversely, not all short-range correlation is dynamic. Moreover, LR dynamic correlation is also an important source of dynamic correlation. See the following section for more details.

6.1.2.1 London dispersion interactions

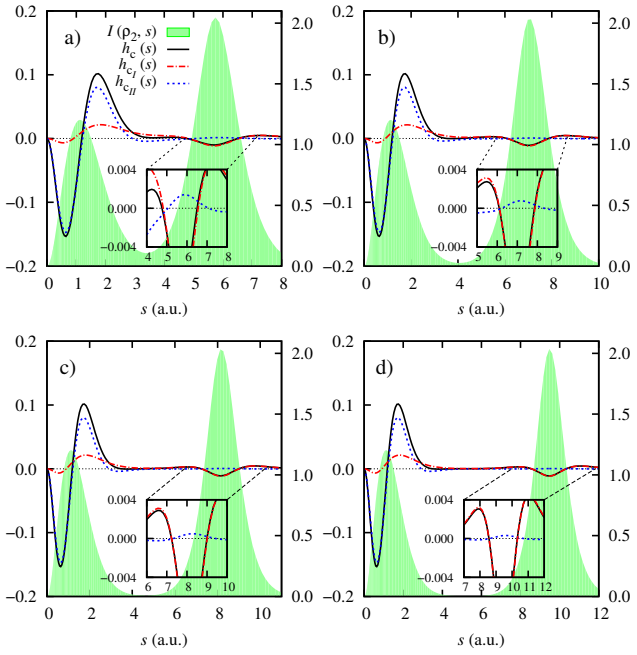


Figure 6.11: Coulomb hole and its correlation components for He_2 at a) 5.67, b) 6.99, c) 8.13, and d) 9.45 a.u. bond distances. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison.

Chapter 1.1.2 introduced the nature of London dispersion forces, where the long-range dynamic correlation of electrons creates a weak bond between two electron clouds. Even though any dimer or molecule experiences dispersion forces due to their ubiquitousness, an easy example of a molecule bounded by dispersion inter-

actions is the helium dimer. Therefore, it is expected to encounter them in the LR region of $h_{c_{II}}$. The Coulomb holes of different geometries of He_2, presented in Fig. 6.11, show a negative minimum in the LR region of h_c depicting intraelectronic interactions. This minimum is mainly defined by h_{c_I} and, apparently, $h_{c_{II}}$ is zero. A closer inspection (see the inset plots in Fig. 6.11) reveals that $h_{c_{II}}$ is not zero but slightly positive, and its area decreases with the stretching of the He–He bond. This reflects the decreasing attraction between both fragments when the interfragment distance R increases, caused by the weakening of dispersion forces. In contrast, the magnitude of h_{c_I} increases (up to a point) with the stretching of a bond, either if the bond is covalent (positive h_{c_I} values) or noncovalent (negative h_{c_I} values). h_{c_I} is also large in systems where HF dissociates into the wrong number of electrons in each fragment. In any of these situations exposed, dispersion interactions can be seen through $h_{c_{II}}$. See the inset plots of Fig. 6.12 for an example of a covalent bond (Be_2), Fig. 6.11 for a noncovalent bond (He_2), and Fig. 6.3 for LiH, where HF dissociates the molecule in two fragments with an uneven number of electrons.

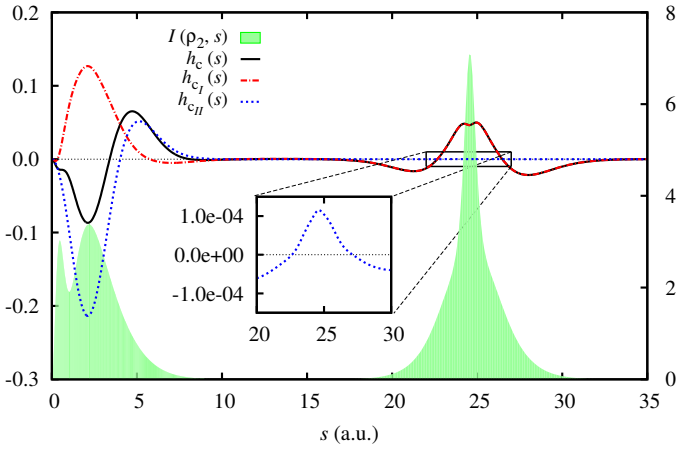


Figure 6.12: Coulomb hole and its correlation components for Be_2 at 24.57 a.u. bond length. The FCI IPD, $I(\rho_2^{\mathrm{FCI}}, s)$ is represented in green (right y-axis) for comparison. The inset plot shows the effect of dispersion interactions in the LR region of $h_{c_{II}}$.

With this, it is demonstrated that $\Delta\rho_2^{c_{II}}$ contains the effects of dispersion interactions, which cannot be seen in the Coulomb hole itself because they are masked by stronger nondynamic, LR interactions. A clear example of that is the beryllium dimer. Although Be_2 is not a van der Waals molecule, dispersion forces are present in any molecule because they are ubiquitous. The inset plot of Fig. 6.12 unveils a small contribution of $h_{c_{II}}$ around the interfragment distance. The values in the LR region of $h_{c_{II}}$ describing dispersion interactions are small since dispersion forces are weak in nature [60], and positive because dispersion forces are binding. Figures

in Chapter 3.2 include more examples for this positive, small maximum in the LR region of $h_{\mathrm{c}_{II}}$ in molecules bound by other forces beside van der Waals.

6.1.3 Conditions for $\Delta\rho_2^{\mathrm{c}_I}$ and $\Delta\rho_2^{\mathrm{c}_{II}}$ and behavior of $h_{\mathbf{c}_I}$ and $h_{\mathbf{c}_{II}}$

The LR asymptotic properties and the coalescense points of $\Delta\rho_2^{\mathrm{c}_I}$ and $\Delta\rho_2^{\mathrm{c}_{II}}$ explain the profiles of h_{c_I} and $h_{\mathrm{c}_{II}}$ in the SR and LR regions of Coulomb holes seen in Sections 6.1.1 and 6.1.2. Concerning the long-range behavior of the 2-PDs, both the HF and exact 2-PD vanish when the interfragment distance increases ($R \to \infty$). On the other hand, the first term of the SD 2-PD (Eq. 1.42) vanishes at large R values because it is a product of electron densities. The second term involves the long-range asymptotics of the 1-rDM, which becomes $\sqrt{\rho(\mathbf{1})}\sqrt{\rho(\mathbf{2})}$ when electron pairs are separated infinitely from each other, and far from any nucleus, as demonstrated by March and Pucci [228]. This is true for systems where its corresponding $(N-1)$–particle system is non-degenerate [229]. The product of electron densities also becomes zero under this condition, and guarantees that both h_{c_I} and $h_{\mathrm{c}_{II}}$ vanish at the large interelectronic distances.

The coalescense points of the c_I and c_{II} hole components are partly determined by the Pauli principle, their corresponding same-spin component being zero when two electrons of the same spin are on top of each other, $r_1 - r_2 = 0$. However, there exists a probability of two electrons with opposite spin being on top of each other. For the c_I component, the condition is determined by Eq. 1.38, the opposite-spin components of the SD 2-PD. This leads to

$$\Delta\rho_2^{\mathrm{c}_I}(\boldsymbol{r_1}, \boldsymbol{r_1}) = 2\left(\rho^{\alpha}(\boldsymbol{r_1})\rho^{\beta}(\boldsymbol{r_1}) - \rho^{\mathrm{HF},\alpha}(\boldsymbol{r_1})\rho^{\mathrm{HF},\beta}(\boldsymbol{r_1})\right). \tag{6.8}$$

As discussed in Section 6.1.1, h_{c_I} is usually positive through all the interelectronic vector axis (with some exceptions, as the LiH molecule presented in Fig. 6.3), including the SR region. This is due to $\Delta\rho_2^{\mathrm{c}_I}(\boldsymbol{r_1}, \boldsymbol{r_1})$ being mainly positive, caused by the HF underestimation of the electron-nucleus cusp. However, as mentioned in Chapter 6.1.1, h_{c_I} is mainly a long-ranged quantity that defines the Coulomb hole at such distances. The LR asymptotic properties described in the former paragraph indicate that this quantity will indeed be positive (again, excepting cases like LiH), since the difference of the product of densities will always be larger.

In the literature, SR interactions or SR correlation is usually used indiscriminately as a synonym for dynamic correlation. Although some LR correlation can

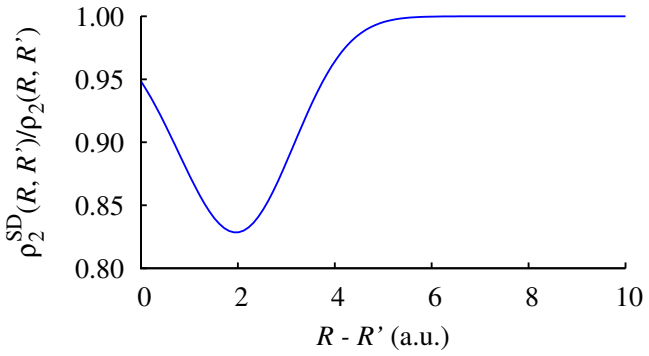


Figure 6.13: $\rho_2^{\mathrm{SD}}(\boldsymbol{R}, \boldsymbol{R}')/\rho_2(\boldsymbol{R}, \boldsymbol{R}')$ ratio against the interfragment distance $R - R'$ for a minimal basis set calculation of H_2.

be dynamical, as London dispersion interactions, it is true that normally dynamic correlation arises from close-electron pair interactions. Through all the molecules studied, it has been spotted that the SR region of $h_{c_{II}}$ always has a Coulomb hole-like profile. It is even present in systems where nondynamic correlation is notorious, with an important contribution of h_{c_I} in the SR region of the hole. $\Delta\rho_2^{c_{II}}$ is able to account for the universality of dynamic correlation at the SR part of the Coulomb hole of any system considered. This can be explained through the on-top probability of the c_{II} component, which reads

$$\Delta\rho_2^{c_{II}}(\boldsymbol{r_1}, \boldsymbol{r_1}) = 2\left(\rho_2^{\alpha\beta}(\boldsymbol{r_1}, \boldsymbol{r_1}) - \rho^{\alpha}(\boldsymbol{r_1})\rho^{\beta}(\boldsymbol{r_1})\right). \tag{6.9}$$

This quantity is usually negative at points close to the nuclei, which are the points that mostly contribute in this coalescence situation, and causes $h_{c_{II}}$ to be negative at short ranges. An analysis through the hydrogen molecule described with a minimal basis set also permits to illustrate this fact (the full development of this issue is given in Chapter 3.2). Taking the leading term of the expansion of $\Delta\rho_2^{c_{II}}(\mathbf{1}, \mathbf{2})$ around the two electron-nucleus cusps, $\Delta\rho_2^{c_{II}}(\boldsymbol{R}, \boldsymbol{R}') = \rho_2(\boldsymbol{R}, \boldsymbol{R}') - \rho_2^{\mathrm{SD}}(\boldsymbol{R}, \boldsymbol{R}')$, it is easy to demonstrate, by means of the representation of $\rho_2^{\mathrm{SD}}(\boldsymbol{R}, \boldsymbol{R}')/\rho_2(\boldsymbol{R}, \boldsymbol{R}')$ against the interatomic separation $R - R'$, that 1) the exact 2-PD is larger than the SD 2-PD at small interatomic separations, and 2) the 2-PD ratio becomes one as the bond is stretched (see Fig. 6.13). This demonstrates that $h_{c_{II}}$ is generally a negative, short-ranged quantity.

6.1.4 Analysis on the spin components of $\Delta\rho_2^{c_I}$ and $\Delta\rho_2^{c_{II}}$

In this section, Coulomb holes and their c_I and c_{II} components of atoms with different multiplicities are examined. By just considering the Coulomb hole, the ground

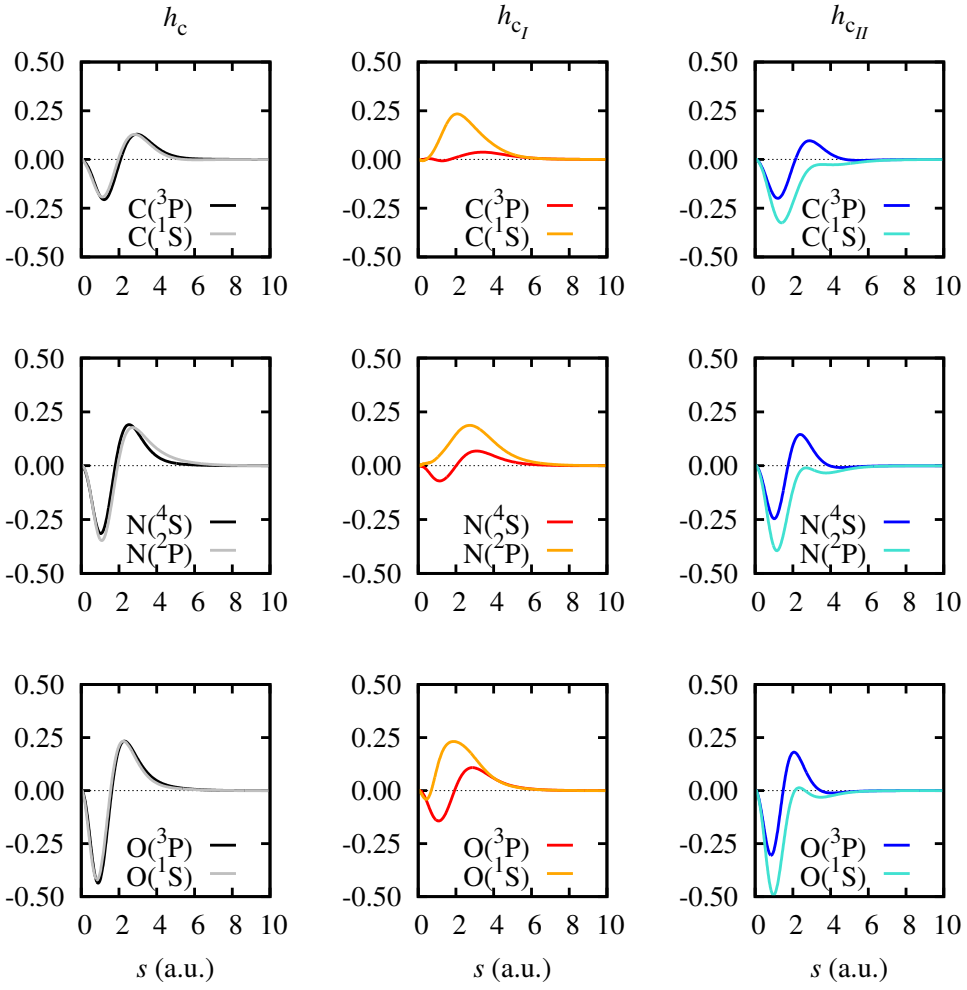


Figure 6.14: Coulomb holes (left column), h_{c_I} (middle column) and $h_{c_{II}}$ (right column) of carbon (top row), nitrogen (middle row) and oxygen (bottom row) atoms in their ground state (darker colors) and minimum multiplicity state (lighter colors).

and minimum multiplicity states of carbon, nitrogen, and oxygen present barely indistinguishable profiles, which may indicate that electron correlation effects are the same for the atom regardless its multiplicity. As well, according to Eq. 1.72, the correlated part of V_{ee} in both multiplicities is the same. Instead, h_{c_I} and $h_{c_{II}}$ reveal different profiles for both multiplicities (see Fig. 6.14). More precisely, similar profiles are generated in the ground state of the three atoms, as well as for the minimum multiplicity cases (see the left column plots in Fig. 6.14). In the ground state cases, both h_{c_I} and $h_{c_{II}}$ present the usual hole shape (negative at the SR region, and positive in the LR region). Instead, for minimum multiplicity cases, h_{c_I} is always positive and $h_{c_{II}}$ is always negative. An explanation for these profiles is attained via a spin-decomposition analysis of the holes, supported by the ability of the IPD to separate into spin components (Eq. 1.73). Fig. 6.15 describes the splitting of the

three holes analyzed into spin components for the carbon atom. The main difference between both multiplicities is found in the $\alpha\alpha$ components of h_{c}, h_{c_I}, and $h_{\mathrm{c}_{II}}$. Conversely, the $\alpha\beta$ holes show similar profiles for both multiplicities.

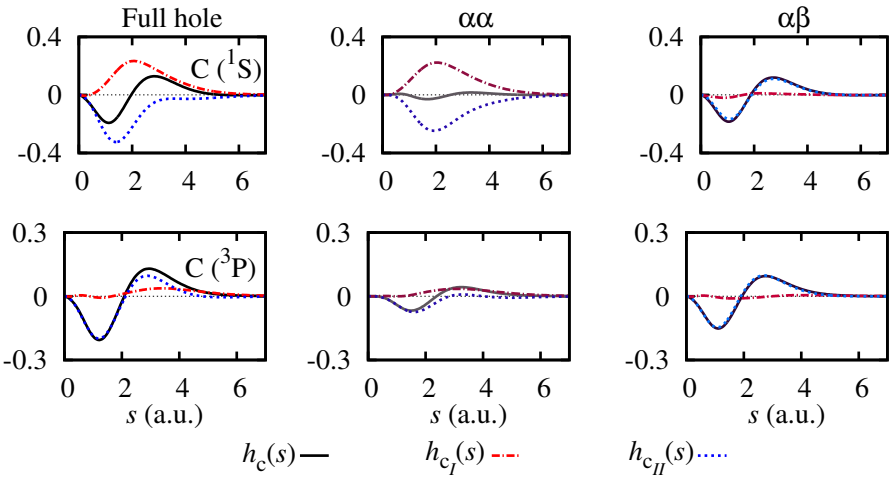


Figure 6.15: Spin components of h_{c}, h_{c_I} and $h_{\mathrm{c}_{II}}$ of singlet (top row) and triplet (bottom row) carbon. (left column) Complete holes (central column) $\alpha\alpha$ component of h_{c}, h_{c_I}, and $h_{\mathrm{c}_{II}}$ (right column) $\alpha\beta$ component of h_{c}, h_{c_I}, and $h_{\mathrm{c}_{II}}$.

The profiles in Fig. 6.15 can be interpreted via an analysis on the spin interactions within the carbon atom. There is a larger amount of $\alpha\alpha$ electron pairs in C(^{3}P), which causes more $\alpha\alpha$ interactions in the atom. Therefore, the 2-PD of triplet carbon contains more $\alpha\alpha$ elements than the singlet. Because of this, the $\alpha\alpha$ component of h_{c} is larger in the triplet state than in the singlet. The SD 2-PD, however, causes a counter effect: the $\alpha\alpha$ component of the HF and SD 2-PD in singlet carbon are, indeed, small. However, the SD 2-PD presents a set of extra $\alpha\alpha$ terms (see Eq. 1.43) which causes it to significantly differ from the $\alpha\alpha$ component of the HF 2-PD. Hence, $\Delta\rho_2^{c_I,\alpha\alpha}$ becomes a large quantity in the singlet state of carbon, and consequently $h_{\mathrm{c}_I}^{\alpha\alpha}$ in nonzero. Instead, the extra $\alpha\alpha$ terms in the SD 2-PD in C(^{3}P) is small compared to the rest of same- and opposite-spin elements in the SD and HF 2-PD and, then, are not enough to cause large differences in $\Delta\rho_2^{c_I,\alpha\alpha}$. Because of this, $h_{\mathrm{c}_I}^{\alpha\alpha}$ is smaller in the triplet. The traces of each 2-PD component are collected in Table 6.2, where $\rho_2^{\alpha\alpha,\mathrm{SD}}$ is the responsible component for the trace deviation in the SD 2-PD.

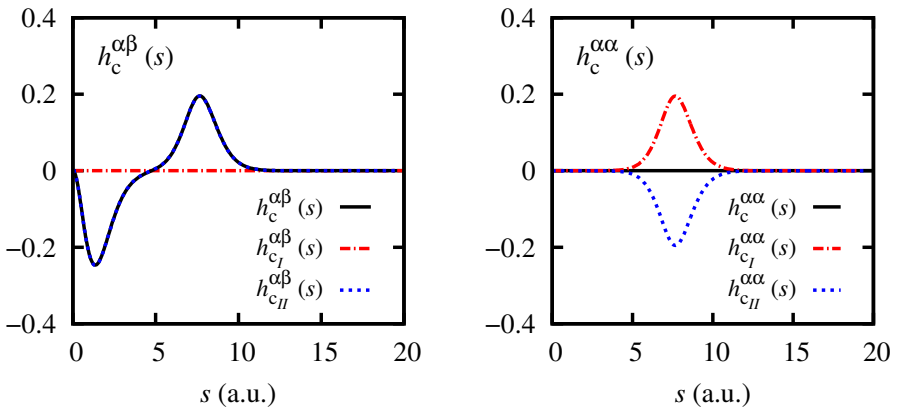


Figure 6.16: The $\alpha\beta$ (left) and $\alpha\alpha$ (right) spin components of the Coulomb hole and the c_I and c_{II} hole components of H_2 with $R = 7.56$ a.u. bond distance.

	$\mathrm{Tr}[\rho_2^{\alpha\alpha}(\mathbf{1},\mathbf{2})]$	$\mathrm{Tr}\left[\rho_2^{\alpha\alpha,\mathrm{SD}}(\mathbf{1},\mathbf{2})\right]$	$\mathrm{Tr}\left[\rho_2^{\alpha\beta}(\mathbf{1},\mathbf{2})\right]$	$\mathrm{Tr}\left[\rho_2^{\alpha\beta,\mathrm{SD}}(\mathbf{1},\mathbf{2})\right]$
Ne	20.0	20.055	25.0	25.0
C(^{3}P)	7.0	7.106	8.0	8.0
C(^{1}S)	6.0	6.587	9.0	9.0
H_2	0.0	1.0	2.0	2.0

Table 6.2: FCI and SD 2-PD traces of the same-spin ($\alpha\alpha$) and opposite-spin ($\alpha\beta$) components of neon, triplet carbon, singlet carbon, and hydrogen molecule in minimal basis. $\mathrm{Tr}[\rho_2^{\alpha\alpha}(\mathbf{1},\mathbf{2})] = N^\alpha(N^\alpha - 1)$ and $\mathrm{Tr}\left[\rho_2^{\alpha\beta}(\mathbf{1},\mathbf{2})\right] = N^\alpha N^\beta$. For open-shell cases, note that the 2-PD containing same-spin interactions must be splitted into all the possible same-spin components, namely the $\alpha\alpha$ and $\beta\beta$ interactions, and the trace must be calculated separately, $\mathrm{Tr}[\rho_2^{\alpha\alpha}(\mathbf{1},\mathbf{2})] = \frac{N^\alpha(N^\alpha-1)}{2}$ and $\mathrm{Tr}\left[\rho_2^{\beta\beta}(\mathbf{1},\mathbf{2})\right] = \frac{N^\beta(N^\beta-1)}{2}$. The trace of $\rho_2^{\alpha\beta,\mathrm{SD}}(\mathbf{1},\mathbf{2})$ is $N^\alpha N^\beta$ by construction.

The minimal-basis hydrogen molecule (Fig. 6.16) also provides an interesting insight on the spin components of the 2-PDs and their respective holes. The profiles obtained are equivalent to the ones described for the atoms discussed in the former paragraph. Because singlet H_2 is only composed of one α and one β electron, the $\alpha\alpha$ component of h_c is practically flat, being the $\alpha\beta$ component responsible for the shape of the total hole, as in the case of C(^{2}S). $h_{c_I}^{\alpha\alpha}$ determines the shape of h_{c_I}. As addressed in the case of C(^{1}S), this profile is due to the extra $\alpha\alpha$ terms in the SD 2-PD which, particularly in singlet H_2, are not present in the HF 2-PD (see Table 6.2). Therefore, the analysis of different multiplicities in a given electronic system provides valuable information on electron correlation according to the distribution of electrons in the molecular orbitals. The shape of the Coulomb hole by itself does

not provide any information on the different multiplicities, because the reference and the HF IPDs are almost identical in both spin states. The partition of the Coulomb hole reveals a different structure and particular profiles according to the multiplicity of the system, granted by the extra terms present in the SD 2-PD.

6.2 On the universal footprint of dispersion interactions in the IPD

Through this thesis, the IPD has been the central tool for analyzing the correlation components of the Coulomb hole. Directly linked with the results discussed in Section 6.1.2.1, where the area of the LR region of $h_{c_{II}}$ decreases with the interatomic distance R because of London dispersion forces, the IPD has been shown to attain a relation with R in analogy to the dispersion energy, ($E^{\text{disp}} \propto R^{-6}$).

An analytical treatment with perturbation theory is performed to the IPD, assuming a model formed by two hydrogenoid atoms separated an infinite distance $R \to \infty$, and treated with a minimal basis composed of Gaussian functions. The zeroth–order wavefunction is

$$\Psi^{(0)}(\mathbf{1},\mathbf{2}) = N\psi_A(\mathbf{1})\psi_B(\mathbf{2}), \tag{6.10}$$

where the superscript (0) labels the zeroth–order correction, A and B refer to the two nuclei, ψ_A is the Gaussian function centered in nucleus A, and N is the corresponding normalization factor of the wavefunction. The perturbation Hamiltonian $\widehat{H}^{(1)}$ contains all the possible LR interactions between nuclei and electrons,

$$\widehat{H}^{(1)} = \frac{1}{R} + \frac{1}{r_{12}} - \frac{1}{r_{A2}} - \frac{1}{r_{B1}}, \tag{6.11}$$

and the Unsöld approximation is applied to obtain the first–order correction to the wavefunction [230],

$$\Psi^{(1)}(\mathbf{1},\mathbf{2}) = \Psi^{(0)}(\mathbf{1},\mathbf{2})\widehat{H}^{(1)}. \tag{6.12}$$

Within these conditions, one can obtain the energy dependence on R for the model system:

$$E^{(2)}(R) = -\frac{6}{\alpha^4 R^6} + \mathcal{O}(R^{-8}), \tag{6.13}$$

being consistent with the widely known dispersion energy, the leading term of which is R^{-6}. Because the zeroth–order correction to the energy is the sum of orbital energies ϵ and the first–order correction is the HF energy, the correction for dispersion interactions is labeled with (2).

The 2-PD of a two-electron system is the square of the wavefunction (consider Eq. 1.16). Its zeroth–order correction is the product of the zeroth–order wavefunction, Eq. 6.10, and the first–order correction is the product of the zeroth– and first–order wavefunction, $\rho_2^{(1)}(\mathbf{1},\mathbf{2}) = 2\Psi^{(0)}(\mathbf{1},\mathbf{2})\Psi^{(1)}(\mathbf{1},\mathbf{2})$ because of the Unsöld approximation described in Eq. 6.12. The zeroth–order correction to the IPD is obtained after

some mathematical manipulation. While the IPD is a function of the interelectronic distance s, we were interested in the specific point where the interelectronic distance vector is equal to the interfragment distance, $s = R$. The zeroth–order correction of $I(R)$ is:

$$I^{(0)}(\rho_2^{(0)}, R) = \left(\frac{\alpha}{16\pi^3}\right)^{1/2} \left(1 - \exp(-4\alpha R^2)\right), \tag{6.14}$$

and, since we want to know the decay at the infinity limit, $R \to \infty$, $I^{(0)}(\rho_2^{(0)}, R)$ yields a constant value corresponding to the distribution of two independent, noninteracting electrons. On the other hand, the first–order correction to the IPD results in a more convoluted expression, but its limit in $R \to \infty$ yields:

$$\lim_{R\to\infty} I^{(1)}(\rho_2^{(1)}, R) = -\frac{4\left(1 + 8\sqrt{2}\right)\alpha^{5/2}}{\pi^{7/2}R^3}, \tag{6.15}$$

where the leading term is R^{-3} because the exponential term in $I^{(0)}$ does not show up in the first–order correction. Therefore, the dependency is a feature of $I^{(1)}$. Since the zeroth–order correction presents a higher-order dependency than the first–order one, the leading term of the IPD (corrected to the first order, $I^1(R)$) is exponential, being $I^1(R) = I^{(0)}(R) + I^{(1)}(R)$. This is the first time that a condition to reproduce dispersion interactions in terms of the pair density is determined.

To unfold the polynomial decay, one must dismiss the exponential leading term of the zeroth–order IPD correction. To do so, the zeroth–order correction has to be substracted from the exact IPD. The HF IPD cannot be considered as $I^{(0)}$ because, since it is constructed with the HF densities, its LR behavior with R is not correct (Eq. 1.37). Instead, the SD IPD is constructed with the exact electron densities (Eq. 1.42), and, consequently, its LR behavior with s correctly reproduces the exact (FCI) decay, $\rho_2(\boldsymbol{r_1}, \boldsymbol{r_2}) \to \rho(\boldsymbol{r_1})\rho(\boldsymbol{r_2})$ (see Fig. 6.17). Hence, $I^{\mathrm{FCI}}(R) - I^{\mathrm{SD}}(R)$ unveils the R^{-3} dependency. For that reason, the SD 2-PD is the minimal 2-PD that guarantees the correct behavior in the LR region of the IPD. The perturbation theory treatment is cut up to the first order because the R^{-3} behavior already shows up at the first-order correction. Since subtracting the SD IPD from the FCI one already recovers the R^{-3} dependency, it demonstrates that higher-order corrections to the IPD do not contain terms with higher order than R^{-3}.

The IPD used to unmask the R^{-3} term coincides with the definition for the hole component responsible for describing dynamic correlation effects, $h_{c_{II}}$. The FCI 2-PD describes all the correlation effects in a system, and, therefore, dispersion interactions are masked by stronger interactions, leading to an exponential leading term, as discussed above. On the other hand, the SD 2-PD can describe some of the

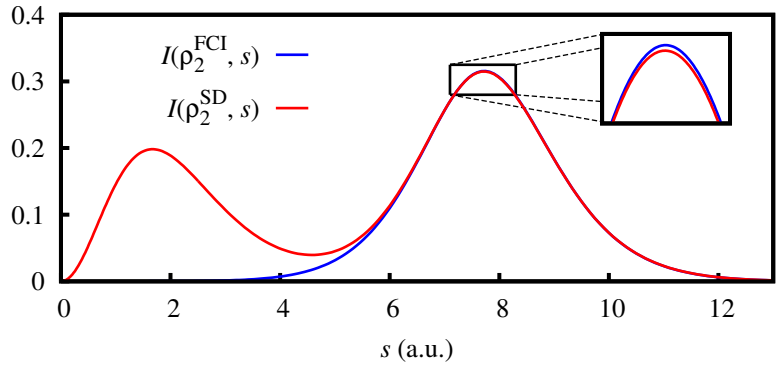
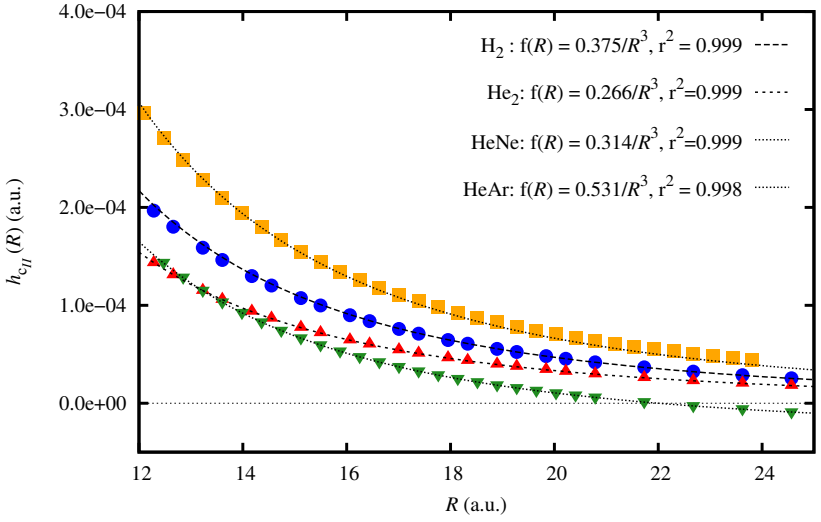


Figure 6.17: The IPDs of the SD 2-PD (red) and the FCI 2-PD (blue) for H_2 at a stretched geometry. The inserted figure is the zoomed section delimited by the square.

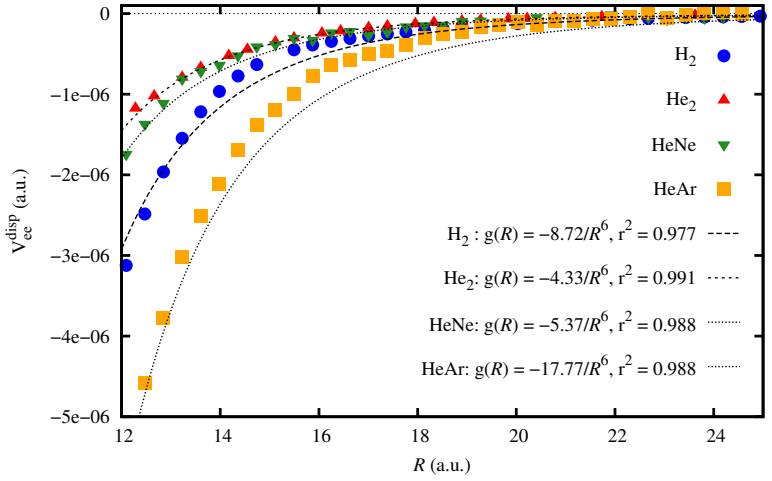
LR interactions and, because of this, it is able to reproduce the LR region of the FCI IPD. This can be seen in the inset plot of Fig. 6.17. After the analysis done to recover the dispersion decay in the IPD, it is sensible to affirm that the SD 2-PD cannot describe LR dynamic correlation effects and, hence, it can only describe LR nondynamic correlation effects. In a sense, this also justifies using the SD 2-PD as an object to partition the 2-PD (Eq. 6.1). Actually, the representation of $h_{c_{II}}(R)$ versus R does follow a R^{-3} decay (figure not shown, consider Fig. 6.18).

Because V_{ee} can be obtained through the IPD via Eq. 1.72, it may seem straightforward to obtain a direct relation with R. However, it is not the case, because V_{ee} is obtained after the full integration of the complete IPD – that is, not only considering $I(R)$, but $I(s)$ being $s \in [0, \infty)$. To our knowledge, there is no explicit dependence of $I(s)$ on R and, hence, an expression for V_{ee}^{disp} is still to be derived. Nevertheless, being V_{ee}^{disp} an energetic term, the dispersion energy decay (R^{-6}, Eq. 1.2) should be retrieved. By considering the correlation decomposition, V_{ee}^{disp} can be obtained as $V_{ee}^{\text{disp}} = V_{ee}^{\text{FCI}} - V_{ee}^{\text{SD}}$ and neglecting the V_{ee} arising from SR interactions. Another way of obtaining V_{ee} arising from LR interactions is by integrating the IPD with Eq. 1.72 using different integration limits. From this procedure, we confirm the V_{ee}^{disp} decay being R^{-6}. Fig. 6.18 contains the V_{ee}^{disp} and the $I^{(1)}(R) \equiv I(\Delta\rho_2^{c_{II}}, R) \equiv h_{c_{II}}(R)$ decays against the interfragment distance R for different van der Waals dimers.

The inclusion of dispersion interactions in quantum chemistry computations requires basis set that includes polarization functions. A system only described by s-type functions (no polarization functions) does not reproduce the R^{-3} decay, and the $h_{c_{II}}(R)$ values (as well as the LR region of the IPD) of both FCI and SD are identical. Whereas an accurate description of such interactions requires a large and



(a)



(b)

Figure 6.18: Compilation and adaptation of the graphs presented in Chapter 3 for the dispersion intracules and energies for H_2 (blue dots), He_2 (red up-pointing triangles), HeNe[1] (green down-pointing triangles) and HeAr (orange squares). The fittings per each component correspond to f(R)=a/R^3 and g(R) = b/R^6. (a) The intracule of the cumulant (which is the definition of $h_{c_{II}}(s)$) evaluated at $s = R$ for several interfragment distances R; and (b) the dispersion V_{ee} component.

flexible basis set, introducing p-type functions to a minimal basis already introduces a minimal description of dispersion interactions in the system [231].

To end this discussion, it is important to recall that most of the dispersion corrections used in the literature are applied *ad hoc*, meaning that they are just energy

[1]The IPD decay ($h_{c_{II}}(R)$ vs. R) in HeNe does not tend to zero but to a negative value, since the angular grids in the present implementation of the RHO2_OPS code (which is used to calculate the IPDs) are limited and do not provide the required precision.

corrections that are not optimized self-consistently. The signature of dispersion presented in this thesis is a quality directly related to the 2-PD, and provides a new means to model such behavior in approximations (for instance, in an exchange-correlation functional in the KS DFT framework), and permits its optimization within the self-consistent field procedure. Note that this modeling is numerically more robust than the usual energetic decay due to the higher power obtained for the IPD condition has, which makes it less prone to numerical errors. Further work is done in this direction by members of Dr. Eduard Matito's group, as well as other researchers that have started to investigate further conditions related to dispersion interactions [114].

6.3 Benchmarking of reduced density matrix approximations

The rDMFT benchmark study presented in Chapter 5 evaluates the physical grounds and predictive abilities of a set of thirteen 2-DM approximations by means of a battery of tests described in the aforementioned chapter. The functionals are evaluated using a set of diatomic molecules composed of first– and second–row elements, and molecules formed by more than two atoms. Two different geometries are contemplated per each diatomic molecule, the equilibrium one, R_{eq}, and a geometry five times larger the equilibrium one, $5R_{\text{eq}}$. For the dissociation energy test, the energy of a geometry fifty times the equilibrium one $50R_{\text{eq}}$ has also been calculated. The set of diatomic molecules includes both open- and closed-shell molecules. As introduced in Section 1.2.3.3, rDMFAs are built in the NOrb representation using the FCI NOccs. These orbitals and occupancies are not optimized, so the benchmark only considers the functional–driven error.

6.3.1 Elements of the density matrix

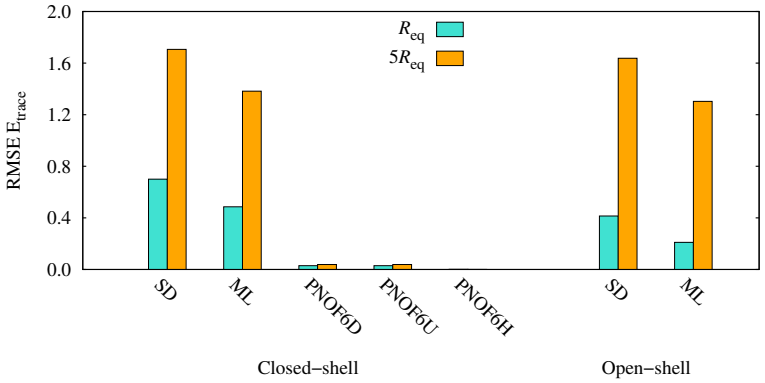


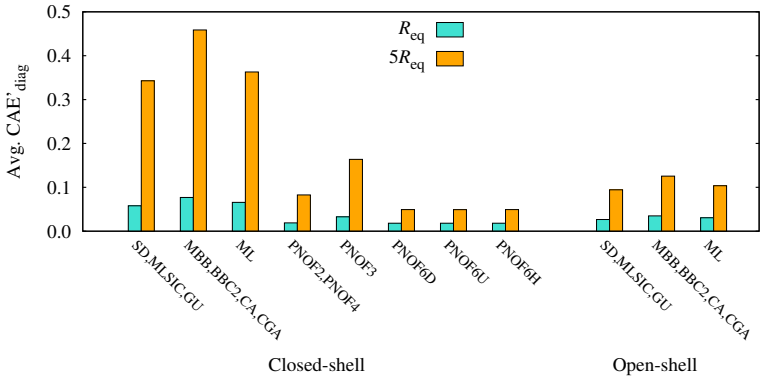
Figure 6.19: RMSE of the trace error (Eq. 22 in Chapter 5), produced by the SD, GU, MLSIC, and PNOF3 group of rDMFAs, labeled under "SD"; the ML approximation, and the three PNOF6 variants. PNOF6h and the other rDMFAs bear the correct trace and are not represented in this graph.

The trace test permits to separate the rDMFAs into two different groups, the first one being composed of approximations that were designed to fulfill the sum rule and bear the correct trace (MBB, BBC2, CA, CGA, PNOF4, and PNOF6h); and a second group that includes the rest of functionals with an incorrect trace, with SD, GU, MLSIC, ML, PNOF3, PNOF6D, and PNOF6u. It is interesting to notice that the traces of the PNOF6d and PNOF6u 2-DMs bear small deviations,

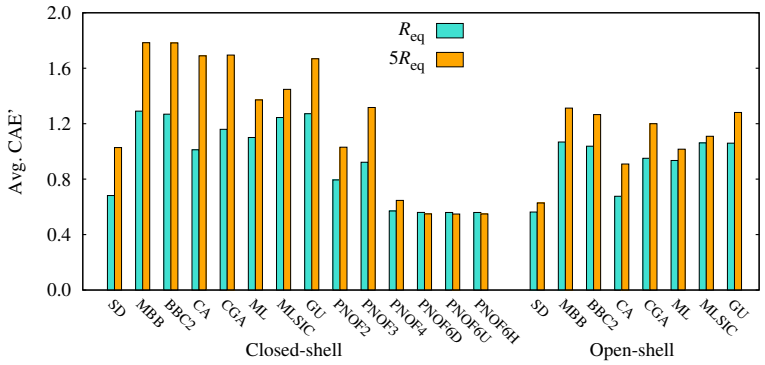
but their combination to produce PNOF6h leads to the exact trace. In addition, SD, GU, MLSIC, and PNOF3 present the same trace error due to the construction of the diagonal elements, and have the largest trace deviation, with ML bearing a close error. If one analyzes the diagonal elements that produce this trace, it is surprising that the rDMFAs that have the correct trace are formed by diagonal elements that are far from coincident with the reference ones (see Fig. 6.20). Most importantly, MBB, BBC2, CA, and CGA show the largest cumulative absolute errors (CAEs) in the diagonal elements, which reveals that more attention is put into fulfilling the sum rule rather than reproducing the correct diagonal elements. On the other hand, the diagonal elements of PNOF2, PNOF4, and PNOF6h are, even though not equal to the reference ones, quite accurately approximated, with a small CAE. The PNOF2 and PNOF4 diagonal elements are exact for systems with two electrons, since they reduce to the exact functional for a two-electron system [232]. Regarding the construction of the complete 2-DM, elements are better reproduced by SD than by the K-functionals, the former not being a rDMFA strictly speaking. PNOF2 and PNOF3 present the same CAE magnitude as SD. PNOF4 and the three versions of PNOF6 produce the most accurate 2-DM elements overall. It is also noticeable that the PNOF family presents an improvement through the versions.

As introduced in Section 1.2.3.2, the N-representability and symmetry conditions assure that a 2-DM belongs to a physical fermionic wavefunction, and guarantees that the obtained energy is variational. The positivity conditions, composed of the P, Q and G matrices, are a known set of N-representability conditions. These conditions are imposed to the PNOF approximations by construction, and, because of this, the PNOF family produce the smallest deviations in this test. PNOF6d is the only rDMFA studied in this thesis that bears a 2-DM that fulfills the three positivity conditions. Instead, PNOF6u and, consequently, PNOF6h generate a very small sum of negative eigenvalues in the P and Q conditions. PNOF3 is the approximation with the largest sum of negative eigenvalues from the PNOF group. PNOFs are designed to produce antysimmetric 2-DMs, and consequently the antisymmetry error is zero for all of these rDMFAs. The CAE of all the elements in the PNOF 2-DMs is smaller than in the K-functionals. With this, the construction of the PNOF family of approximations imposes some of the known physical requirements that the exact 2-DM fulfills. They produce the smallest amount of negative eigenvalues from the positivity conditions, produce fermionic 2-DMs, and provide the closest 2-DM elements to the reference 2-DM.

The SD approximation fulfills the P and G conditions, but the amount of negative



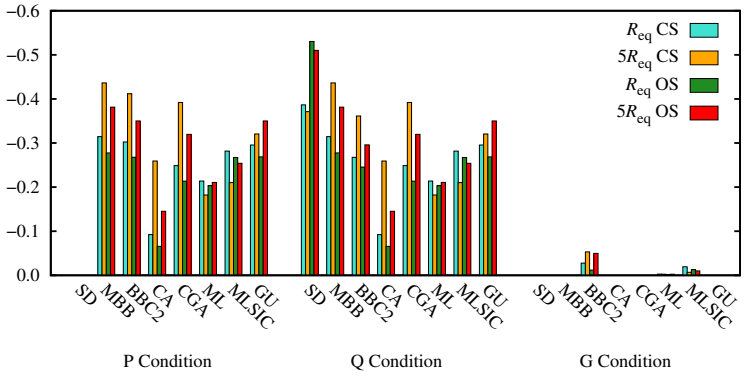
(a)



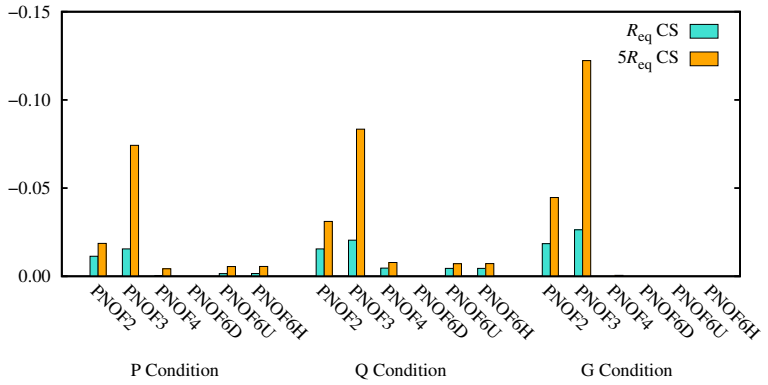
(b)

Figure 6.20: (a) Average of the cumulative absolute errors (CAEs) produced by the diagonal elements of the approximate 2-DM (Eq. 26 from Chapter 5). (b) Average of the CAEs of all the 2-DM elements (Eq. 27 from Chapter 5).

eigenvalues for the Q condition is excessively large. Instead, CA generates the smallest sum of negative eigenvalues in both P and Q conditions, and fulfills the G condition. CA is, then, the K-functional that better accomplishes the known positivity conditions. The G condition is satisfied by all the K-functionals, with the exception of BBC2, MLSIC, and ML. However, the amount of negative eigenvalues in the other two positivity conditions is large. Some K-functionals bear the correct trace by imposition, but the diagonal elements that also contribute in the trace value deviate from the reference ones. In general, only the G-condition is attained by the K-only approximations, but a large sum of negative eigenvalues are obtained for the other two conditions. The CAE of the 2-DM elements differ from the reference 2-DM, and the antisymmetry error is different from zero for all the K-functionals.



(a)



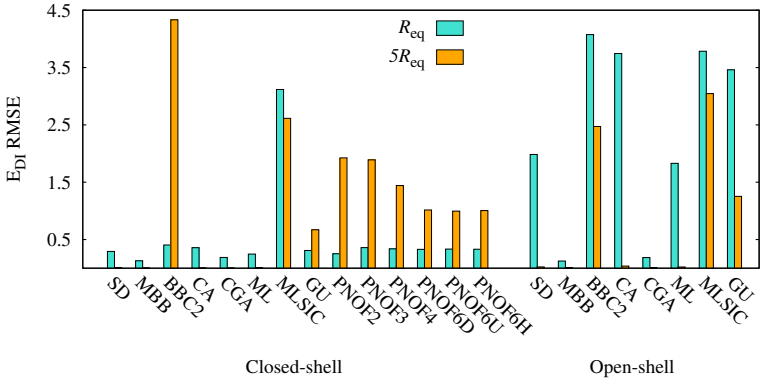
(b)

Figure 6.21: (a) Average sum of negative eigenvalues for the P, Q and G matrices of the (top) K-functionals (b) PNOF family for the open-shell (OS) and closed-shell (CS) set of molecules.

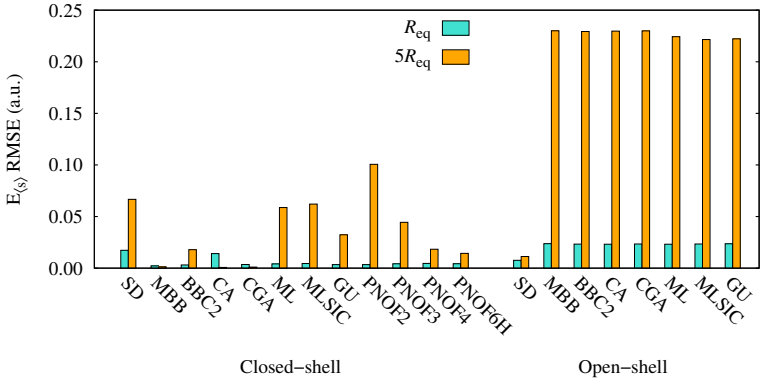
6.3.2 Delocalization indexes and electron pair distributions

The electron pair distribution of the approximate 2-DMs is analyzed considering the delocalization index (DI) predictions and the intracule probability density (IPD) profiles. The delocalization index permits to obtain a measure of the covalent bond order between two fragments (see Chapter 5 for more details). Usually, the DI is a positive value, yet it can take negative numbers in some particular cases. However, the molecule set used in this study presents no situation in which the DI can be negative. MLSIC, GU, BBC2, and PNOF3 produce some negative DIs; in fact, MLSIC and GU only predicted one positive DI for closed-shell molecules at stretched geometries. This leads to the conclusion that such versions do not perform any better than their corresponding self-interaction uncorrected versions (ML and MBB, respectively). Because of this, these functionals sometimes present underestimated DI values, and some of the largest root mean squared errors (RMSEs) in the DI test (see Fig. 6.22). Regarding the PNOF family, they produce larger errors in the DI

prediction than the K-functionals. PNOF2 shares the same error magnitude with PNOF3, caused by erroneous predictions in the DI of stretched H_2. This is also connected with the erroneous IPD obtained for that molecule, where PNOF2 produces larger probabilities at the SR region. However, the main result obtained from the DI test is that PNOFs are not able to correctly distribute electron pairs in dissociated molecules. When a molecule is dissociated, its bond order and, therefore, its DI are zero. This is generally predicted correctly by the K-functionals, and thus their error is smaller at $5R_{\text{eq}}$ than at R_{eq}. However, all the members of the PNOF family produce nonzero DI values at these stretched geometries. This inability to distribute the electron pairs is connected to their performance in the IPD test: PNOFs give lower electron pair probabilities at long ranges than FCI, indicating that electrons are still at the bonding region.



(a)



(b)

Figure 6.22: RMSE of (a) the error committed in reproducing the delocalization index DI (Eq. 29 in Chapter 5), and (b) the relative error in predicting the average interelectronic distance (Eq. 33 in Chapter 5), for closed- and open-shell molecules.

In general, PNOF4 and PNOF6 present good IPDs in molecules with a small

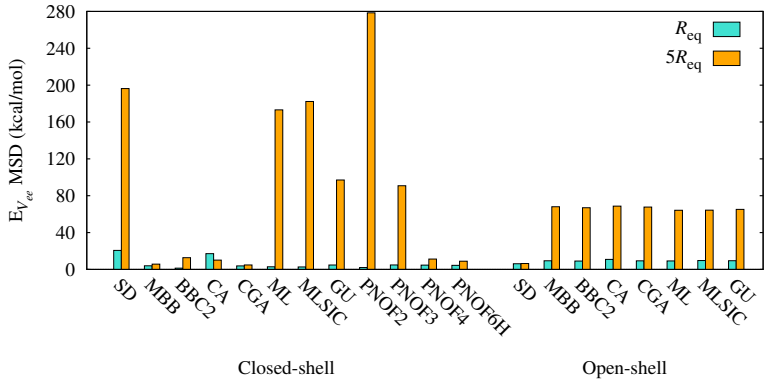
number of electrons (6 electrons) when found at the equilibrium geometry (see Figs. S1–S4 in Chapter 5). This is in agreement with the closing remarks obtained for the harmonium atom [136], in which these PNOFs present satisfactory IPDs. The PNOF4 and PNOF6 IPDs are very similar through the systems studied, most of the times undistinguishable one from the other. Their behavior in molecules with more than 6 electrons is analogous to the SD profile, but with a smaller magnitude, showing that electron correlation is slightly underestimated by these rDMFAs (see, for instance, Fig. S10 in Chapter 5). As stated, they present a higher probability at medium ranges of the IPD, which causes $E_{I(s)}$ (Eq. 31 in Chapter 5) to be largely negative at mid ranges and positive at the LR region. PNOF3 generates large negative IPD values (negative probability) in stretched H_2, which is completely incorrect. Still at stretched geometries, PNOF3 usually places an excessive amount of electron pairs in the LR region of the IPD, whereas PNOF2 behaves similarly to PNOF4 and PNOF6 at those regions, but with larger $E_{I(s)}$ magnitudes (see the LR region of Fig. S31 in Chapter 5 as an example).

In the set of closed-shell molecules, MBB, the first rDMFA ever designed, presents the smallest deviations of the DI predictions, the expected value of the interelectronic distance ($\langle s \rangle$) and its variance (σ^2), and, in general, does not show large deviations from the reference IPD. It also provides good DI values for open-shell molecules, but this may be serendipitous because its IPD is not a good approximation (*vide infra*). The behavior of CGA in these tests is similar to MBB, and, on the contrary, ML, MLSIC, and GU (along with PNOF3), show the largest errors in these tests (recall that some of these approximations produce negative DI values). Whereas BBC2 produces really good approximations of $\langle s \rangle$ and its variance, it also generates incorrect equilibrium DI predictions, and the BBC2 $E_{I(s)}$ is moderately acceptable. CA presents wiggly $E_{I(s)}$ profiles, indicating several flaws in most of the electron pair ranges of the IPD. However, the CA DI, $\langle s \rangle$ and σ^2 predictions represent one of the smallest errors (only for closed-shell molecules).

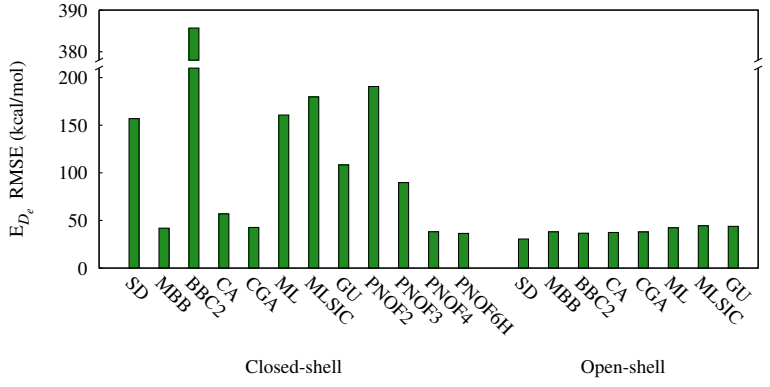
Open-shell molecules present different error trends than the set of closed-shell molecules. Errors for the DI and the IPD are generally too large for open-shell molecules. Even though CGA and MBB present quite correct DI predictions for this set of molecules, the IPD profiles are not so accurate. SD always provides an IPD description closer to FCI than the rest of K-functionals, all of which generate highly similar IPD. This implies that all of them share the same error for open-shell cases. This is also seen in the $\langle s \rangle$ test, where all the K-functionals produce approximately the same error. In some molecules, $\langle s \rangle$ is largely overestimated. The

comparison of the IPD of singlet and triplet O_2 shows that rDMFAs produce an IPD for the triplet state which is very alike to the FCI IPD of the singlet state; in other words, the rDMFAs seem to approximate the molecule as if it was a closed-shell system.

All in all, the oldest and simplest rDMFA is the one that generates the best approximate 2-DM to reproduce the electron pair distribution in the diatomic molecules studied. Whereas PNOFs evolved correctly through their versions, their improvement is not enough for describing stretched geometries. It would be interesting to evaluate PNOF5 and PNOF7, which are not considered in this work due to their perfect pairing approach, to verify whether the whole PNOF family presents a problem in describing DI for stretched molecules. The self-interaction corrected functionals of ML and MBB (MLSIC and GU) do not present any better results than the uncorrected versions (where MBB is the best-performing rDMFA).



(a)



(b)

Figure 6.23: (a) RMSE of the relative error in the electronic repulsion potential (Eq. 37 in Chapter 5) in kcal/mol. (b) RMSE of the error in the dissociation repusion potential (Eq. 38 in Chapter 5) in kcal/mol.

6.3.3 Energies

One of the main outcomes obtained from the benchmarking study is that rDMFAs reproduce V_{ee} reasonably well for equilibrium geometries, but their performance becomes particularly poor for dissociated molecules. Another remarkable point is that SD produces smaller errors in the set of open-shell molecules, once again confirming that rDMFAs are not suited for multiplicities other than a singlet. However, the error in open-shell molecules is smaller than the error generated at geometries other than the equilibrium. As a consequence, excessively large deviations are committed in dissociation energy (D_e) predictions: the smallest RMSE obtained in the dissociation energy error is about 50 kcal/mol, which is far from the chemical accuracy required for computational studies. Instead, errors at equilibrium geometries display a maximum RMSE of 20 kcal/mol, produced by SD. The rest of rDMFAs (excepting CA) generate a RMSE below 5 kcal/mol in the closed-shell set of molecules, and an average error of 10 kcal/mol for open-shell molecules. All in all, the most appropriate rDMFAs to calculate electronic repulsion energies are MBB and CGA, since they provide the smallest RMSE at both R_{eq} and $5R_{\text{eq}}$, and thus produce the smallest RMSE in the D_e test. Other qualitatively good rDMFAs are CA, PNOF4, and PNOF6, as they also show rather small errors with a consistent magnitude for the three geometries.

6.3.3.1 Size-extensivity of the rDMFA

The ability to correctly predict V_{ee} with the size of a molecule is tested with a systematic increase of hydrogen atoms (and a systematic increase of electron repulsion) in the H_N polyhedral molecules already introduced in Chapter 3.1 of this thesis, being N the number of hydrogen atoms, placed at the vertices of the corresponding polynomial shape. Hydrogens are separated 10 Å from the center of mass, so very mild dynamic correlation effects are present in the system. To evaluate the size extensivity, the energetic predictions are divided by the number of hydrogens (Eq. 40 in Chapter 5).

All the K-only RDMFAs excepting BBC2 and GU give size-extensive results, or show small deviations frm size-extensivity. This does not mean, however, that their predictions are accurate. Similarly to the outcome obtained in the V_{ee} test of diatomic molecules, only MBB, CA, and CGA produce an almost 0 error, whereas the $E_{V_{ee}/N}$ in SD, ML, and MLSIC is around 100 kcal/mol. The GU error is smaller than this latter group of approximations, yet neither as good as MBB, CA, and CGA, nor size-extensive. Instead, BBC2 produces excessively large errors and predicts

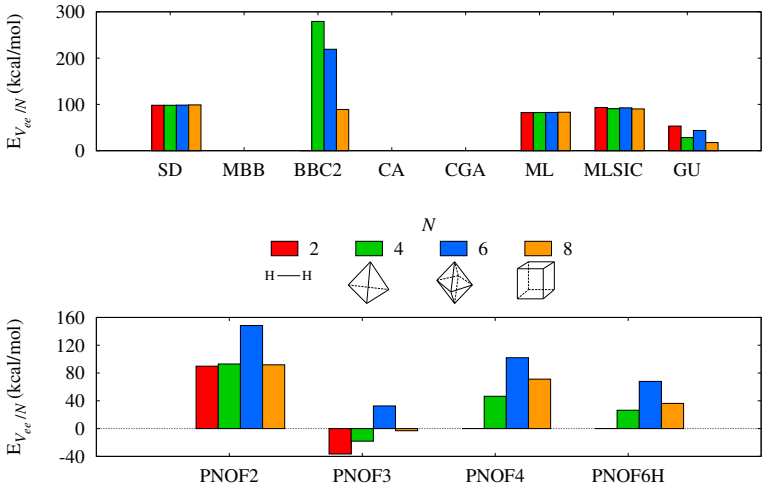


Figure 6.24: Error in the interelectronic repulsion energy divided by the number of hydrogen atoms (Eq. 40 in Chapter 5), for N hydrogen atoms placed at the vertices of an N-vertex polyhedron, in which the H atoms are separated by 10 Å from the geometrical center.

larger repulsion in H_6 than in H_8. On the other hand, no member of the PNOF family seems to be size-extensive. PNOF3 predicts an unphysical, negative V_{ee} in H_2, quite in line with the underestimated V_{ee} values obtained from the V_{ee} test. PNOF4 and PNOF6 produce very reliable energy values for H_2, in agreement with the V_{ee} test as well, yet their predictions are not as accurate in the rest of H_N molecules. Being PNOF2 the version with the largest $E_{V_{ee}/N}$, the other PNOFs show qualitatively lower errors.

6.3.3.2 Description of nondynamic correlation

Also introduced in Chapter 3.1, the D_{2h}/D_{4h} potential energy surface of H_4 permits to tune the amount of nondynamic correlation in the molecule by changing the angle formed between two adjacent hydrogen atoms and the geometrical center, θ, and the distance between a hydrogen atom from the geometrical center, R. As indicated in the article in Chapter 5, we consider the error produced in predicting the energetic barrier or path going from the equilibrium geometry ($\theta = 70°$ and $R = 0.8$ Å [233]) to other three geometries (remaining combinatinos of $\theta = 90°$ or $\theta = 70°$, and $R =$ 0.8 Å or $R = 4.0$ Å). For simplicity, the transition from D_{2h} to D_{4h} that keeps the same distance between hydrogen atoms R is named path 1 (red bars in Fig. 6.25), the path that keeps the same symmetry but involves the stretching of the bond is referred as path 2 (green bars in Fig. 6.25) and the barrier that changes both the symmetry and the distance between hydrogens is path 3 (blue bars in Fig. 6.25).

Path 1 is the best approximated by the rDMFAs. The transition brings short-range nondynamic correlation in the system, proceeding from an orbital degeneracy. Even though the errors in this path are smaller than in paths 2 and 3, the average error committed is around 100 kcal/mol, which is far from the chemical accuracy. The rDMFAs with errors below 65 kcal/mol are MBB, BBC2, CA, CGA, PNOF2, and PNOF6h, yet the magnitude of the deviation is larger than 50 kcal/mol. Errors are generally three times larger for the other two paths, in which the stretching of the H–H bond is involved. Similar errors are obtained for both paths, indicating that most of the rDMFAs evaluated in this work present difficulties in describing nondynamic correlation effects arising from molecular dissociations. This is also in agreement with the results obtained in the V_{ee} test, where the smallest errors are produced when evaluating equilibrium geometries.

The smallest errors obtained in this test are compressed between 10 and 15 kcal/mol, and are produced by MBB and CGA for paths 2 and 3. Instead, BBC2, MLSIC, and PNOF2 generate larger errors than SD. PNOF2 seems to be able to describe nondynamic correlation coming from short-range orbital degeneracies (small deviation in path 1). PNOF3 predicts the barrier energy of path 2 quite accurately, but the deviations in paths 1 and 3 are huge. GU and MLSIC produce larger deviations than their self-interaction uncorrected versions, MBB and ML, again showing that such correction does not bring any improvement to the rDMFA in the tests carried out in this thesis. On the other hand, CGA does present a better ability to include nondynamic correlation than CA.

The PNOF group are not the best performing rDMFAs in this test, and contradicts the conclusions obtained in a former study carried out in our laboratory [233], where PNOF6 produced qualitatively correct energies and orbitals for H_4. In contrast to this work, the geometries, occupation numbers, and orbitals were optimized using PNOF6h. However, the respective absolute energies for the equilibrium geometry produce relative errors of the same order of magnitude as the ones obtained in this work. In the former study, PNOF6h committed consistent deviations through the geometries analyzed, whereas the results obtained in this test show a drastic error increase for the path involving both the symmetry transition and the H–H bond stretching (blue bars in Fig. 6.25). This suggests that the functional–driven error is larger than the 1-DM–driven error in PNOF6h.

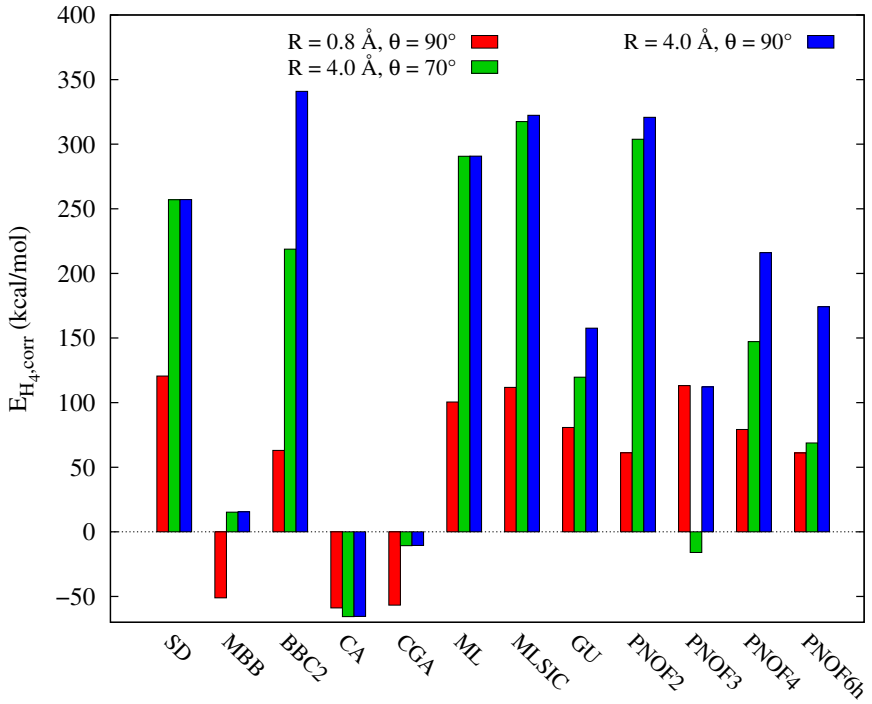


Figure 6.25: Error of the D_{2h}/D_{4h} potential energy curve of H_4, $E_{H_4,\mathrm{corr}} = (V^{\mathrm{X}}_{ee,\ \mathrm{geom}} - V^{\mathrm{X}}_{ee,\ \mathrm{ref}}) - (V^{\mathrm{FCI}}_{ee,\ \mathrm{geom}} - V^{\mathrm{FCI}}_{ee,\ \mathrm{ref}})$ (Eq. 41 in Chapter 5), where the repulsion energies are calculated with respect to the interelectronic potential of the ground state geometry, $R = 0.8$ Å and $\theta = 70°$.

6.3.4 Description of dispersion interactions

To study the ability of the rDMFAs to include dispersion forces in their description, we take advantage of the signature of London dispersion interactions introduced in Chapter 4 and discussed in Section 6.2. $\mathrm{E}_{\mathrm{disp}}(R) = \left(I(\rho_2^{\mathrm{X}}, R) - I(\rho_2^{\mathrm{SD}})\right) R^3$ (Eq. 42 in Chapter 5) assesses the IPD difference with respect to the SD approximation (see the discussion in Section 6.2) and multiplies it by R^3. In this way, it permits easy identification of the power dependence of each rDMFA with R. The IPDs of the hydrogen molecule and the helium van der Waals dimer are the systems chosen for this test. If the rDMFA correctly portrays dispersion interactions, $\mathrm{E}_{\mathrm{disp}}(R)$ is positive and constant along R (within the dissociation regime). The approximations that have an exponential or a lower power decay with R exhibit an $\mathrm{E}_{\mathrm{disp}}(R)$ that rapidly goes to zero at early values of R. In contrast, approximations with a larger power dependency grow to large (absolute) values. Fig. 6.26 contains the analysis for H_2 and He_2 molecules. In H_2, PNOF2, PNOF3, GU, and MLSIC present the latter type of behavior introduced (larger power dependency than R^{-3}), whereas, in the rest of rDMFAs, $\mathrm{E}_{\mathrm{disp}}(R)$ rapidly tends to zero. In addition to PNOF2, PNOF3, GU, and MLSIC, BBC2 and the rest of PNOFs also present a growing $\mathrm{E}_{\mathrm{disp}}(R)$ trend in He_2. This indicates that the PNOF family of functionals overestimate dispersion forces,

as they describe it with a larger dependency than R^{-3}. ML, CGA, CA, and MBB present smaller power dependencies.

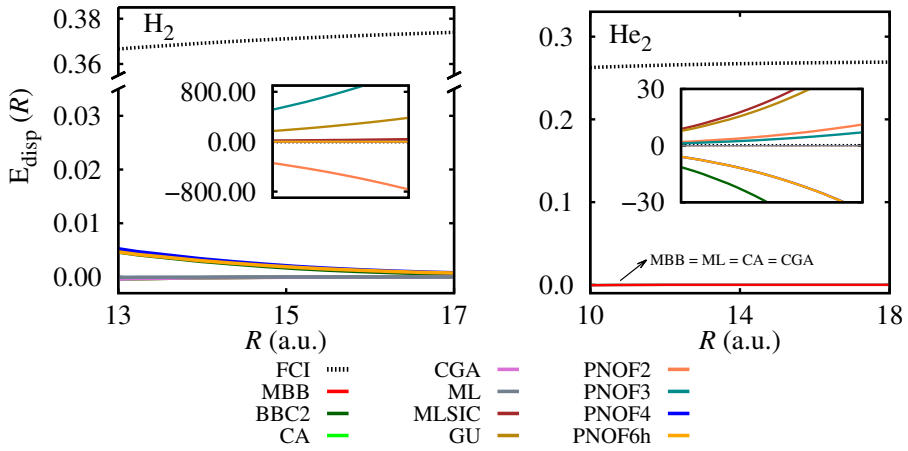


Figure 6.26: $E_{disp}(R)$ profiles (Eq. 42 in Chapter 5) of (left) the hydrogen molecule and (right) the helium van der Waals dimer, according to each rDMFA. The dashed, black line corresponds to the reference (FCI) behavior.[2] The inset plot in both graphs contains a zoomed-out version of the graph (*i.e.* a greater y-axis scale). In graph (b), the PNOF4 line is under the PNOF6h one.

In a former study, it was concluded that PNOF2 presents the ability to bind two He atoms to form the van der Waals dimer [161,200]. However, it seems that PNOF2 is able to bind the helium dimer due to an incorrect assessment of dispersion interactions. Using the FCI NOccs and NOrbs, we have calculated the interelectronic repulsion for different geometries, and the Coulomb hole for the He_2 dimer at 9.45 a.u. of bond distance, to compare with the results presented by Piris and coworkers [161, 200], in which the occupancies and orbitsls were optimized at the PNOF2 level. The Coulomb hole and V_{ee} values are in agreement with the published results, and it apparently seems that PNOF2 describes dispersion forces correctly. We have also computed the $h_{c_{II}}$ component of the Coulomb hole, which compares the FCI/PNOF2 IPD with the SD IPD. Analysis of $h_{c_{II}}$ indicates that the PNOF2 electron-pair probability is excessively large at the LR region of the Coulomb hole (see Fig. 6.27). Additionally, according to the results obtained with $E_{disp}(R)$, this LR probability increases with R, when it should decrease. Therefore, the performance of PNOF2 in the dispersion test suggests that the higher-power dependency with R gives PNOF2 the ability to bind He_2, caused by an incorrect treatment of

[2]Note that the FCI E_{disp} (R) function (dashed line) in H_2 should be constant or decreasing (if the R^{-4} and further power dependencies were important) with R, but it is not. The angular grids in the present implementation of the RHO2_OPS code (used to integrate the IPDs) are limited and do not provide the required precision.

dispersion forces.

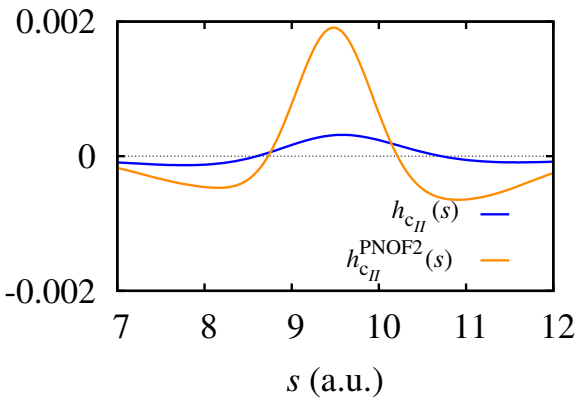


Figure 6.27: The LR region of the $h_{c_{II}}(s) = I(\rho_2^{FCI}, s) - I(\rho_2^{SD}, s)$ component of the Coulomb hole (blue line) and $h_{c_{II}}^{PNOF2}(s) = I(\rho_2^{PNOF2}, s) - I(\rho_2^{SD}, s)$ (orange line) for the helium dimer at 9.45 a.u. of bond distance.

The IPD dependence that describes dispersion interactions has the form of $\left(I^{FCI} - I^{SD}\right) \propto 1/R^3$, being a positive quantity that becomes zero as R increases. In H_2, only BBC2, PNOF4, and PNOF6 present positive $E_{disp}(R)$ values, and would be the rDMFAs that produce the closest description to dispersion interactions, yet it has been shown that they are unable to do so in He_2. ML, CGA, CA, and MBB produce negative $E_{disp}(R)$. PNOF2 produces negative $E_{disp}(R)$ values for H_2, but at He_2 $E_{disp}(R)$ are positive, and the contrary behavior is produced by BBC2, PNOF4, and PNOF6. GU, MLSIC, and PNOF3 $E_{disp}(R)$ values are positive in both molecules. Therefore, no rDMFA is able to reproduce the universal signature of dispersion interactions in the IPD, and therefore the description of London dispersion forces is not correctly described by any of the rDMFAs. Whereas some functionals produce smaller power dependencies, others bind molecules too strongly. rDMFAs with larger power dependencies may bind a van der Waals dimer, even though the treatment of the interactions is not correct, as the PNOF2 description for the helium dimer [161].

Chapter 7

Conclusions

This thesis introduces a separation of the pair density into two correlation components in terms of range separation. The versatility of the IPD permits the further separation of the Coulomb hole into two hole components, which we have named $h_{\mathrm{c}_I}(s)$ and $h_{\mathrm{c}_{II}}(s)$. The two pair density components that define the hole components, $\Delta\rho_2^{\mathrm{c}_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{\mathrm{c}_{II}}(\mathbf{1},\mathbf{2})$, are able to capture long- and short-range interactions, as well as nondynamic and dynamic correlation effects.

Physical models and small molecules have been used to test and validate the correlation decomposition scheme and perform a range separation study of electron correlation. The short-ranged interactions arisen from close electron pairs give rise to dynamic correlation, which is, therefore, universal in molecules with (at least) two electrons. The c_{II} hole component has shown to be never small through all the molecules studied, either if the dominant correlation in the molecule is dynamic or not. Generally, $h_{\mathrm{c}_{II}}(s)$ defines the shape of the short-range part of the Coulomb hole. Systems with orbital degeneracy have an important contribution of short-range nondynamic correlation, which is correctly accounted for by the c_I contribution of the hole. In these cases, the short-range region of the Coulomb hole is described by both c_I and c_{II} components. Systems composed of a minimum of two fragments present a small positive area of $h_{\mathrm{c}_{II}}(s)$ in the long-range region of the Coulomb hole for large interfragment distances. This positive area becomes smaller when the interfragment distance is increased. In fact, the evaluation of $h_{\mathrm{c}_{II}}(R)$ at different geometries reveals a decay of $h_{\mathrm{c}_{II}}(R) \propto 1/R^3$, which arises from the universal footprint of dispersion interactions introduced in this thesis. The c_{II} component is, therefore, able to describe interactions arisen from London dispersion forces. $h_{\mathrm{c}_{II}}(s)$ does not define the long-range profile of the Coulomb hole, since dispersion forces are weak; instead, $h_{\mathrm{c}_I}(s)$ is the dominant component in this region, and can be positive (in covalent bonds) or negative (in non-covalent bonds, and in cases where Hartree–Fock does

not dissociate into fragments with an even number of electrons). The c_I and c_{II} correlation components produce different profiles in atoms with two different multiplicities. Whereas the Coulomb holes of two different spin states of an atom are undistinguishable, $h_{c_I}(s)$ and $h_{c_{II}}(s)$ are clearly different in both cases. Moreover, $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ is able to discern between types A and B nondynamic correlation. To sum up this point, c_{II} is generally important for describing short-range interactions, and $h_{c_{II}}(s)$ is always present in the short-range part of the Coulomb hole. On the other hand, c_I is predominant at long interelectronic ranges, mostly arisen from the orbital degeneracies occurring when a molecule dissociates. By analyzing systems dominated by either dynamic or nondynamic correlation, we have identified that the c_I component is small in systems mainly described by dynamic correlation, and c_I is dominant when there is an important contribution of nondynamic correlation. Therefore, the first two points of the first objective of this thesis are accomplished: $\Delta\rho_2^{c_I}(\mathbf{1},\mathbf{2})$ and $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ are able to separate the correlation pair density $\Delta\rho_2^{c}(\mathbf{1},\mathbf{2})$ and the Coulomb hole $h_c(s)$ into two components dominated by short- and long-range interactions.

A universal signature of dispersion interactions has been found in the IPD of the cumulant of the 2-PD. As stated, the cumulant is coincident with the c_{II} correlation component, which mainly contains the dynamic correlation effects of the 2-PD. When the IPD is evaluated at the point where the interelectronic distance equals with the interatomic separation $s = R$, a relation with R similar to the dispersion energy behavior is obtained, $I\left(\Delta\rho_2^{c_{II}}, R\right) \propto 1/R^3$. This trend is numerically more robust than the energetic one, $E^{\text{disp}} \propto 1/R^6$. This concludes the third point of the first objective of this thesis. Note that, because the IPD is a functional of the 2-PD, this signature is explicitly dependent on the 2-PD. This opens the door in new imprelemtations for treating London dispersion interactions, as they can be treated in a fully self-consistent manner if the long-range decay is imposed in further approximate methods.

The second objective of this thesis was to provide a comprehensive benchmarking of reduced density matrix functional approximations. The study illustrates some of the strengths and flaws of some of the rDMFAs available nowadays. The inability to properly describe open-shell molecules and stretched geometries indicates that much work is still to be done in the field. No rDMFA is able to correctly describe dispersion interactions, as none of them provides an $1/R^3$ decay in their IPD. Whereas some approximations have been said to be dispersion-including methods [161, 234], it has been demonstrated that they produce binding potential energy curves due to

an incorrect treatment of dispersion forces.

The MBB approximation has provided the best performances and predictions through all the benchmarking tests. Being the first rDMFA ever designed, it is able to produce accurate electron-electron repulsive potentials and delocalization indices for any geometry and multiplicity considered. The approximate V_{ee} are size-extensive for the H_N set of molecules, and bears one of the smallest errors in systems dominated by nondynamic correlation. Its dissociation energies present one of the smallest deviations as well. The MBB 2-DM is not purely antisymmetric and does not fulfill de positivity conditions, caused by the simple approximation in the exchange elements. CA and CGA have also shown outstanding results in the benchmarking tests, in line with the performance of MBB. CGA shows some improvement with respect to its predecessor, with smaller deviations in V_{ee} predictions, as well as in dissociation energies, energies in cases where nondynamic correlation is present, and intracule probability densities. CA has a smaller antisymmetry error, and it better fulfills de positivity conditions than CGA. Whereas the CA IPD profiles are the most irregular among the other rDMFAs studied, CGA presents IPDs similar to MBB. Results of the BBC2 approximation have been outstanding in some tests, but not so good in others. The BBC2 IPD behavior is in line with MBB and CGA IPDs, showing really small errors in $\langle s \rangle$ and its variance. The V_{ee} predictions for the diatomic set are quite accurate. However, the V_{ee} predictions are not size-extensive for the H_N set, even predicting larger repulsion in H_6 than in H_8. It provides an excessively huge error in the D_e test, an incorrect description of the D_{2h}/D_{4h} potential energy surface, and erroneous DI predictions. rDMFAs with empirical parameters do not outperform other rDMFAs, suggesting that parameterization may not be sufficient to produce accurate 2-DMs. The self-interaction corrected versions of MBB and ML do not perform qualitatively better than their original versions either. The construction of PNOF approximations involves the imposition of several physical conditions, such as the antisymmetry of the 2-DM, and the positivity conditions to the 2-DM. Because of this, they perform better than the K-functionals in the N-representability test, and produce no error in the antisymmetry test. The 2-DM elements are correctly approximated to the reference 2-DM, producing smaller CAEs than the K-functionals. However, this does not guarantee an accurate performance in other tets. It has been seen than PNOFs present problems in describing the bond order in dissociated molecules, and PNOF3 has produced unphysical interelectronic repulsion potentials. PNOF2 and PNOF4 perform outstandingly in two-electron systems because they reduce to the exact functional, yet the extrapolation for larger systems is not so satisfactory for PNOF2. PNOF4 and the three versions of PNOF6

are the best performing versions among the PNOF family. The error in their dissociation energy predictions is equivalent to the one produced by MBB. Their V_{ee} predictions are, however, not size-extensive, and produce non-negligible errors in systems with nondynamic correlation. It would be interesting an analysis using the newest PNOF versions.

In general, there is room for improvement for all the rDMFAs, and a study of the 1-DM–driven error would be advantageous for the community, with complementary conclusions to the benchmarking work presented in this thesis. Moreover, the tests in which the SD approximation has provided poor results provides fruitful information about the correlation effects that $\Delta\rho_2^{c_{II}}(\mathbf{1},\mathbf{2})$ must contain.

Bibliography

[1] S. Wilson, *Electron correlation in molecules*. New York: Dover, 2007.

[2] W. Kutzelnigg and P. von Herigonte, "Electron Correlation at the Dawn of the 21st Century," in *Advances in Quantum Chemistry* (P.-O. Löwdin, ed.), vol. 36, pp. 185–229, Elsevier Academic Press, 1999.

[3] E. Wigner and F. Seitz, "On the constitution of metallic sodium," *Phys. Rev.*, vol. 43, pp. 804–810, 1933.

[4] E. Wigner and F. Seitz, "On the constitution of metallic sodium. II," *Phys. Rev.*, vol. 46, pp. 509–524, 1934.

[5] J. Slater, "Ferromagnetism and the band theory," *Rev. Mod. Phys.*, vol. 25, no. 1, p. 199, 1953.

[6] P.-O. Löwdin, "Correlation problem in many-electron quantum mechanics i. review of different approaches and discussion of some current ideas," *Adv. Chem. Phys.*, vol. 2, pp. 207–322, 1959.

[7] E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, "Ground-state correlation energies for two- to ten-electron atomic ions," *Phys. Rev. A*, vol. 44, no. 11, p. 7071, 1991.

[8] D. Pines, "Electron interaction in metals," in *Solid State Physics. Advances in Research and Applications* (F. Seitz and D. Turnbull, eds.), vol. 1, pp. 367–450, New York: Academic Press Inc. Publishers, 1955.

[9] J. A. Pople and J. S. Binkley, "Correlation energies for AH_n molecules and cations," *Mol. Phys.*, vol. 29, pp. 599–611, 1975.

[10] E. K. U. Gross, M. Petersilka, and T. Grabo, "Conventional Quantum Chemical correlation energy versus Density-Functional correlation energy," in *Chemical Applications of Density-Functional Theory* (B. B. Laird, R. B. Ross, and T. Ziegler, eds.), vol. 629 of *ACS Symposium Series*, ch. 3, pp. 42–53, Washington, DC: American Chemical Society, 1996.

[11] W. Kutzelnigg, G. Del Re, and G. Berthier, "Correlation coefficients for electronic wave functions," *Phys. Rev.*, vol. 172, no. 1, p. 49, 1968.

[12] J. Cioslowski, "Differential density matrix overlap: an index for assessment of electron correlation in atoms and molecules," *Theor. Chim. Acta (Berlin)*, vol. 81, no. 4, pp. 319–327, 1992.

[13] K. Raghavachari and J. B. Anderson, "Electron correlation effects in molecules," *J. Phys. Chem.*, vol. 100, pp. 12960–12973, 1996.

[14] P. Ziesche, "On relations between correlation, fluctuation and localization," *J. Mol. Struct. (Theochem)*, vol. 527, no. 1-3, pp. 35–50, 2000.

[15] D. Cremer, "Density functional theory: coverage of dynamic and non-dynamic electron correlation effects," *Mol. Phys.*, vol. 99, no. 23, pp. 1899–1940, 2001.

[16] T. Juhász and D. A. Mazziotti, "The cumulant two-particle reduced density matrix as a measure of electron correlation and entanglement," *J. Chem. Phys.*, vol. 125, no. 17, p. 174105, 2006.

[17] D. P. Tew, W. Klopper, and T. Helgaker, "Electron correlation: The many-body problem at the heart of chemistry," *J. Comput. Chem.*, vol. 28, no. 8, pp. 1307–1320, 2007.

[18] S. Grimme, "Density functional theory with London dispersion corrections," *WIREs, Comput. Mol. Sci.*, vol. 1, no. 2, pp. 211–228, 2011.

[19] J. Klimeš and A. Michaelides, "Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory," *J. Chem. Phys.*, vol. 137, no. 12, p. 120901, 2012.

[20] T. Björkman, A. Gulans, A. V. Krasheninnikov, and R. M. Nieminen, "van der Waals bonding in layered compounds from advanced density-functional first-principles calculations," *Phys. Rev. Lett.*, vol. 108, no. 23, p. 235502, 2012.

[21] J. T. Skolnik and D. A. Mazziotti, "Cumulant reduced density matrices as measures of statistical dependence and entanglement between electronic quantum domains with application to photosynthetic light harvesting," *Phys. Rev. A*, vol. 88, no. 3, p. 032517, 2013.

[22] A. Ambrosetti, D. Alfè, R. A. Distasio Jr, and A. Tkatchenko, "Hard numbers for large molecules: Toward exact energetics for supramolecular systems," *J. Phys. Chem. Lett.*, vol. 5, no. 5, pp. 849–855, 2014.

[23] J. Hermann, D. Alfè, and A. Tkatchenko, "Nanoscale π–π stacked molecules are bound by collective charge fluctuations," *Nat. Commun.*, vol. 8, p. 14052, 2017.

[24] J. E. Lennard-Jones, "The spatial correlation of electrons in molecules," *J. Chem. Phys.*, vol. 20, p. 1024, 1952.

[25] M. Buijse and E. Baerends, "Fermi holes and Coulomb holes," in *Density Functional Theory of Molecules, Clusters, and Solids* (D. E. Ellis, ed.), vol. 12, ch. 1, pp. 1–46, Springer Netherlands, 1996.

[26] J. Hirschfelder and J. Linnett, "The energy of interaction between two hydrogen atoms," *J. Chem. Phys.*, vol. 18, no. 1, pp. 130–142, 1950.

[27] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*. Chichester: John Wiley & Sons, Ltd, 2000.

[28] O. Sinanoğlu, "Many-electron theory of atoms, molecules and their interactions," *Adv. Chem. Phys.*, vol. 6, pp. 315–412, 1964.

[29] D. K. W. Mok, R. Neumann, and N. C. Handy, "Dynamic and nondynamic correlation," *J. Phys. Chem.*, vol. 100, pp. 6225–6230, 1996.

[30] G. C. Lie and E. Clementi, "Study of the electronic structure of molecules. XXI. Correlation energy corrections as a functional of the Hartree-Fock density and its application to the hydrides of the second row atoms," *J. Chem. Phys.*, vol. 60, no. 4, pp. 1275–1287, 1974.

[31] H. J. Silverstone and O. Sinanoğlu, "Many-electron theory of nonclosed-shell atoms and molecules. i. orbital wavefunction and perturbation theory," *J. Chem. Phys.*, vol. 44, no. 5, pp. 1899–1907, 1966.

[32] R. J. Bartlett and J. F. Stanton, "Applications of Post-Hartree–Fock methods: A tutorial," *Rev. Comp. Chem.*, vol. 5, pp. 65–169, 1994.

[33] M. Via-Nadal, M. Rodríguez Mayorga, E. Ramos-Cordoba, and E. Matito, "Singling out dynamic and nondynamic correlation," *J. Phys. Chem. Lett.*, vol. 10, no. 14, pp. 4032–4037, 2019.

[34] R. J. Bartlett and M. Musiał, "Coupled-cluster theory in quantum chemistry," *Rev. Mod. Phys.*, vol. 79, no. 1, p. 291, 2007.

[35] J. W. Hollett and P. M. Gill, "The two faces of static correlation," *J. Chem. Phys.*, vol. 134, no. 11, p. 114111, 2011.

[36] T. Kato, "On the eigenfunctions of many-particle systems in quantum mechanics," *Comm. Pure Appl. Math.*, vol. 10, pp. 151–177, 1957.

[37] R. McWeeny, "Present Status of the Correlation Problem," in *The New World of Quantum Chemistry* (B. Pullman and R. Parr, eds.), vol. 2, pp. 3–31, Reidel, Dordrecht: Springer Netherlands, 1976.

[38] F. London, "Zur Theorie und Systematik der Molekularkräfte," *Z. Physik*, vol. 63, no. 3–4, pp. 245–279, 1930.

[39] J. H. van Lenthe, J. G. C. M. Duijneveldt van de Rijdt, and F. B. van Duijneveldt, "Weakly Bonded Systems," in *Ab Initio Methods in Quantum Chemistry Part 2* (K. P. Lawley, ed.), vol. 69, pp. 521–566, New York: Wiley, 1987.

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

[41] S. R. White and D. A. Huse, "Numerical renormalization-group study of low-lying eigenstates of the antiferromagnetic s=1 heisenberg chain," *Phys. Rev. B*, vol. 48, no. 6, p. 3844, 1993.

[42] C. Møller and M. S. Plesset, "Note on an approximation treatment for many-electron systems," *Phys. Rev.*, vol. 46, no. 7, p. 618, 1934.

[43] I. Shavitt, "The method of configuration interaction," in *Methods of electronic structure theory*, pp. 189–275, Springer, 1977.

[44] J. Čížek, "On the correlation problem in atomic and molecular systems. Calculation of wavefunction components in Ursell-type expansion using quantum-field theoretical methods," *J. Phys. Chem.*, vol. 45, no. 11, pp. 4256–4266, 1966.

[45] K. Hirao, ed., *Recent Advances in Multireference Methods*, vol. 4. Singapore: World Scientific Publishing Co. Pte. Ltd., 1999.

[46] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, "Second-order perturbation theory with a casscf reference function," *J. Phys. Chem.*, vol. 94, no. 14, pp. 5483–5488, 1990.

[47] K. Andersson, P.-Å. Malmqvist, and B. O. Roos, "Second-order perturbation theory with a complete active space self-consistent field reference function," *J. Chem. Phys.*, vol. 96, no. 2, pp. 1218–1226, 1992.

[48] R. J. Buenker and S. D. Peyerimhoff, "Individualized configuration selection in ci calculations with subsequent energy extrapolation," *Theor. Chim. Acta (Berlin)*, vol. 35, no. 1, pp. 33–58, 1974.

[49] H.-J. Werner and P. J. Knowles, "An efficient internally contracted multiconfiguration–reference configuration interaction method," *J. Chem. Phys.*, vol. 89, no. 9, pp. 5803–5814, 1988.

[50] D. A. Mazziotti, "Anti-hermitian contracted schrödinger equation: Direct determination of the two-electron reduced density matrices of many-electron molecules," *Phys. Rev. Lett.*, vol. 97, no. 14, p. 143002, 2006.

[51] E. Pastorczak, M. Hapka, L. Veis, and K. Pernal, "Capturing the dynamic correlation for arbitrary spin-symmetry CASSCF reference with adiabatic connection approaches: Insights into the electronic structure of the tetramethyleneethane diradical," *J. Phys. Chem. Lett.*, vol. 10, no. 16, pp. 4668–4674, 2019.

[52] J. W. Hollett and P.-F. Loos, "Capturing static and dynamic correlation with ΔNO-MP2 and ΔNO-CCSD," *J. Chem. Phys.*, vol. 152, no. 1, p. 014101, 2020.

[53] A. Savin, "A combined density functional and configuration interaction method," *Int. J. Quantum Chem.*, vol. 34, no. S22, pp. 59–69, 1988.

[54] S. Grimme and M. Waletzke, "A combination of Kohn–Sham density functional theory and multi-reference configuration interaction methods," *J. Chem. Phys.*, vol. 111, no. 13, pp. 5645–5655, 1999.

[55] M. Piris, "Global method for the electron correlation," *Phys. Rev. Lett.*, vol. 119, p. 063002, 2017.

[56] J. Toulouse, F. Colonna, and A. Savin, "Long-range short-range separation of the electron-electron interaction in density-functional theory," *Phys. Rev. A*, vol. 70, p. 062505, 2004.

[57] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, "A long-range correction scheme for generalized-gradient-approximation exchange functionals," *J. Chem. Phys.*, vol. 115, no. 8, pp. 3540–3544, 2001.

[58] J. Toulouse, I. C. Gerber, G. Jansen, A. Savin, and J. G. Angyán, "Adiabatic-Connection Fluctuation-Dissipation Density-Functional Theory based on range separation," *Phys. Rev. Lett.*, vol. 102, no. 9, p. 096404, 2009.

[59] I. Shavitt, "The treatment of electron correlation. Where do we go from here?," in *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules* (C. E. D. (ed.), ed.), NATO ASI Series 133, pp. 185–196, Springer Netherlands, 1984.

[60] A. J. Stone, *The theory of intermolecular forces*. Oxford University Press, 2nd ed., 2013.

[61] J. N. Israelachvili, *Intermolecular and Surface Forces*. London: Academic Press, 2nd ed., 1991.

[62] J. H. Van der Waals, *On the continuity of the gases and liquid state*. PhD thesis, Doctoral Dissertation, Leiden University, 1873.

[63] R. H. French, V. A. Parsegian, R. Podgornik, R. F. Rajter, A. Jagota, J. Luo, D. Asthagiri, M. K. Chaudhury, Y.-m. Chiang, S. Granick, *et al.*, "Long range interactions in nanoscale science," *Rev. Mod. Phys.*, vol. 82, no. 2, pp. 1887–1944, 2010.

[64] S. Grimme, "Density functional theory with London dispersion corrections," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 1, no. 2, pp. 211–228, 2011.

[65] A. Tkatchenko, "Current understanding of van der Waals effects in realistic materials," *Adv. Funct. Mater.*, vol. 25, no. 13, pp. 2054–2061, 2015.

[66] M. Stöhr, T. Van Voorhis, and A. Tkatchenko, "Theory and practice of modeling van der Waals interactions in electronic-structure calculations," *Chem. Soc. Rev.*, vol. 48, pp. 4119–4154, 2019.

[67] C. A. Coulson, *Valence*. London: Oxford University Press, 2 ed., 1965.

[68] Y. V. Shtogun and L. M. Woods, "Many-body van der Waals interactions between graphitic nanostructures," *J. Phys. Chem. Lett.*, vol. 1, no. 9, pp. 1356–1362, 2010.

[69] F. Rozpłoch, J. Patyk, and J. Stankowski, "Graphenes bonding forces in graphite," *Acta Physica Polonica A*, vol. 112, pp. 557–562, 2007.

[70] R. A. DiStasio, O. A. von Lilienfeld, and A. Tkatchenko, "Collective many-body van der Waals interactions in molecular systems," *Proc. Natl. Acad. Sci. U.S.A*, vol. 109, no. 37, pp. 14791–14795, 2012.

[71] C. I. Branden and J. Tooze, *Introduction to Protein Structure*. New York/London: Garland Science Publishing, 2nd ed., 1999.

[72] D. J. Scheeres, C. M. Hartzell, P. Sánchez, and M. Swift, "Scaling forces to asteroid surfaces: The role of cohesion," *Icarus*, vol. 210, no. 2, pp. 968–984, 2010.

[73] K. Autumn, Y. A. Liang, S. T. Hsieh, W. Zesch, W. P. Chan, T. W. Kenny, R. Fearing, and R. J. Full, "Adhesive force of a single gecko foot-hair," *Nature*, vol. 405, no. 6787, pp. 681–685, 2000.

[74] K. Autumn and A. M. Peattie, "Mechanisms of adhesion in geckos," *Integr. Comp. Biol.*, vol. 42, pp. 1081–1090, 12 2002.

[75] K. Autumn, M. Sitti, Y. A. Liang, A. M. Peattie, W. R. Hansen, S. Sponberg, T. W. Kenny, R. Fearing, J. N. Israelachvili, and R. J. Full, "Evidence for van der Waals adhesion in gecko setae," *Proc. Natl. Acad. Sci. U.S.A*, vol. 99, no. 19, pp. 12252–12256, 2002.

[76] P. Debye, "Molekularkräfte und ihre Elektrische Deutung," *Physik. Z.*, vol. 22, pp. 302–308, 1921.

[77] W. Keesom, "Die van der Waalsschen Kohäsionskräfte," *Physik. Z.*, vol. 22, pp. 129–141, 1921.

[78] L. Pauling and E. B. Wilson, *Introduction to Quantum Mechanics with Applications to Chemistry*. Dover Books on Physics, New York: Dover Publications, Inc., 1985.

[79] H. Margenau and N. R. Kestner, *Theory of Intermolecular Forces*, vol. 18 of *International Series of Monographs in Natural Philosophy*. Oxford: Pergamon Press, 2nd ed., 2013.

[80] F. London, "Über einige Eigenschaften und Anwendungen der Molekularkräfte," *Z. Physik. Chem. B*, vol. 11, pp. 222–251, 1930.

[81] J. E. Lennard-Jones, "On the determination of molecular fields.–II. From the equation of state of gas," *Proc. R. Soc. London Ser. A*, vol. 106, no. 738, pp. 463–477, 1924.

[82] H. B. Casimir, "On the attraction between two perfectly conducting plates," in *Proc. Kon. Ned. Akad. Wet. B*, vol. 51, p. 793, 1948.

[83] H. B. Casimir and D. Polder, "The influence of retardation on the London–van der Waals forces," *Phys. Rev.*, vol. 73, no. 4, pp. 360–372, 1948.

[84] S. K. Lamoreaux, "The Casimir force: background, experiments, and applications," *Rep. Prog. Phys.*, vol. 68, no. 1, pp. 201–236, 2004.

[85] J. Hermann, R. A. DiStasio Jr, and A. Tkatchenko, "First-principles models for van der Waals interactions in molecules and materials: Concepts, theory, and applications," *Chem. Rev.*, vol. 117, no. 6, pp. 4714–4758, 2017.

[86] M. Shahbaz and K. Szalewicz, "Dispersion energy from local polarizability density," *Phys. Rev. Lett.*, vol. 122, p. 213001, 2019.

[87] J. F. Dobson, K. McLennan, A. Rubio, J. Wang, T. Gould, H. M. Le, and B. P. Dinte, "Prediction of dispersion forces: is there a problem?," *Aust. J. of Chem.*, vol. 54, no. 8, pp. 513–527, 2001.

[88] R. G. Parr and Y. Weitao, *Density-Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry, Oxford University Press, 1994.

[89] J. F. Dobson and T. Gould, "Calculation of dispersion energies," *J. Phys.: Condens. Matter*, vol. 24, no. 7, p. 073201, 2012.

[90] G. A. DiLabio and A. Otero-de-la Roza, "Noncovalent interactions in Density Functional Theory," in *Reviews in Computational Chemistry* (A. L. Parrill and K. B. Lipkowitz, eds.), vol. 29, ch. 1, pp. 1–97, New Jersey: John Wiley & Sons, Inc., 2016.

[91] J. Hermann and A. Tkatchenko, "Electronic exchange and correlation in van der Waals systems: Balancing semilocal and nonlocal energy contributions," *J. Chem. Theory Comput.*, vol. 14, pp. 1361–1369, 2018.

[92] G. A. DiLabio, "Atom-centered potentials for noncovalent interactions and other applications," in *Non-Covalent Interactions in Quantum Chemistry and Physics*, ch. 7, pp. 221–240, Elsevier, 2017.

[93] J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, and G. K.-L. Chan, "Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy," *Science*, vol. 345, no. 6197, pp. 640–643, 2014.

[94] F. Furche, "Molecular tests of the random phase approximation to the exchange-correlation energy functional," *Phys. Rev. B*, vol. 64, no. 19, p. 195120, 2001.

[95] D. Lu, Y. Li, D. Rocca, and G. Galli, "Ab initio calculation of van der Waals bonded molecular crystals," *Phys. Rev. Lett.*, vol. 102, no. 20, p. 206411, 2009.

[96] R. Peverati and D. G. Truhlar, "Quest for a universal density functional: the accuracy of density functionals across a broad spectrum of databases in chemistry and physics," *Philos. Trans. R. Soc. A*, vol. 372, no. 2011, p. 20120476, 2014.

[97] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist, "Van der Waals density functional for general geometries," *Phys. Rev. Lett.*, vol. 92, no. 24, p. 246401, 2004.

[98] K. Lee, É. D. Murray, L. Kong, B. I. Lundqvist, and D. C. Langreth, "Higher-accuracy van der Waals density functional," *Phys. Rev. B*, vol. 82, no. 8, p. 081101, 2010.

[99] O. A. Vydrov and T. Van Voorhis, "Nonlocal van der Waals density functional made simple," *Phys. Rev. Lett.*, vol. 103, no. 6, p. 063004, 2009.

[100] O. A. Vydrov and T. Van Voorhis, "Nonlocal van der Waals density functional: The simpler the better," *J. Chem. Phys.*, vol. 133, no. 24, p. 244103, 2010.

[101] A. D. Becke and E. R. Johnson, "A density-functional model of the dispersion interaction," *J. Chem. Phys.*, vol. 123, no. 15, p. 154101, 2005.

[102] E. R. Johnson and A. D. Becke, "Local-hybrid functional based on the correlation length," *J. Chem. Phys.*, vol. 124, p. 174104, 2006.

[103] A. D. Becke and E. R. Johnson, "Exchange-hole dipole moment and the dispersion interaction revisited," *J. Chem. Phys.*, vol. 127, no. 15, p. 154108, 2007.

[104] F. O. Kannemann and A. D. Becke, "van der Waals interactions in Density-Functional Theory: rare-gas diatomics," *J. Chem. Theory Comput.*, vol. 5, no. 4, pp. 719–727, 2009.

[105] F. O. Kannemann and A. D. Becke, "van der Waals interactions in Density-Functional Theory: intermolecular complexes," *J. Chem. Theory Comput.*, vol. 6, no. 4, pp. 1081–1088, 2010.

[106] F. O. Kannemann and A. D. Becke, "Atomic volumes and polarizabilities in Density-Functional Theory," *J. Chem. Phys.*, vol. 136, no. 3, p. 034109, 2012.

[107] A. Tkatchenko and M. Scheffler, "Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data," *Phys. Rev. Lett.*, vol. 102, no. 7, p. 073005, 2009.

[108] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," *J. Chem. Phys.*, vol. 132, no. 15, p. 154104, 2010.

[109] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth, and S. Grimme, "A generally applicable atomic-charge dependent London dispersion correction," *J. Chem. Phys.*, vol. 150, no. 15, p. 154122, 2019.

[110] A. Tkatchenko, R. A. DiStasio Jr, R. Car, and M. Scheffler, "Accurate and efficient method for many-body van der Waals interactions," *Phys. Rev. Lett.*, vol. 108, no. 23, p. 236402, 2012.

[111] A. Ambrosetti, A. M. Reilly, R. A. DiStasio Jr, and A. Tkatchenko, "Long-range correlation energy calculated from coupled atomic response functions," *J. Chem. Phys.*, vol. 140, no. 18, p. 18A508, 2014.

[112] P. L. Silvestrelli, "Van der Waals interactions in density functional theory by combining the quantum harmonic oscillator-model with localized Wannier functions," *J. Chem. Phys.*, vol. 139, no. 5, p. 054106, 2013.

[113] D. P. Kooi and P. Gori-Giorgi, "A variational approach to London dispersion interactions without density distortion," *J. Phys. Chem. Lett.*, vol. 10, no. 7, pp. 1537–1541, 2019.

[114] O. Werba, A. Raeber, K. Head-Marsden, and D. A. Mazziotti, "Signature of van der Waals interactions in the cumulant density matrix," *Phys. Chem. Chem. Phys.*, vol. 21, pp. 23900–23905, 2019.

[115] C. A. Coulson, "Present state of molecular structure calculations," *Rev. Mod. Phys.*, vol. 32, no. 2, pp. 170–177, 1960.

[116] J. Von Neumann, "Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik," *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, vol. 1927, pp. 245–272, 1927.

[117] P. A. M. Dirac, "Note on exchange phenomena in the Thomas atom," *Proc. Camb. Phil. Soc.*, vol. 26, no. 2, pp. 376–385, 1930.

[118] R. McWeeny, "Some recent advances in density matrix theory," *Rev. Mod. Phys.*, vol. 32, no. 2, pp. 335–369, 1960.

[119] E. R. Davidson, *Reduced density matrices in quantum chemistry*, vol. 6. New York: Academic Press, Inc., 1976.

[120] A. J. Coleman and V. I. Yukalov, *Reduced density matrices: Coulson's challenge*, vol. 72. Berlin: Springer Verlag, 2000.

[121] P.-O. 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," *Phys. Rev.*, vol. 97, no. 6, pp. 1474–1489, 1955.

[122] F. Bopp, "Ableitung der Bindungsenergie vonN-Teilchen-Systemen aus 2-Teilchen-Dichtematrizen," *Z. Physik*, vol. 156, no. 3, pp. 348–359, 1959.

[123] A. J. Coleman, "Density matrices in the quantum theory of matter: Energy, intracules and extracules," *Int. J. Quantum Chem.*, vol. 1, no. S1, pp. 457–464, 1967.

[124] D. A. Mazziotti, "Approximate solution for electron correlation through the use of schwinger probes," *Chem. Phys. Lett.*, vol. 289, no. 5, pp. 419–427, 1998.

[125] W. Kutzelnigg and D. Mukherjee, "Cumulant expansion of the reduced density matrices," *J. Chem. Phys.*, vol. 110, no. 6, pp. 2800–2809, 1999.

[126] C. Valdemoro, "Approximating the second-order reduced density matrix in terms of the first-order one," *Phys. Rev. A*, vol. 45, no. 7, p. 4462, 1992.

[127] D. A. Mazziotti, "Contracted schrödinger equation: Determining quantum energies and two-particle density matrices without wave functions," *Phys. Rev. A*, vol. 57, no. 6, pp. 4219–4234, 1998.

[128] K. Ruedenberg, "The physical nature of the chemical bond," *Rev. Mod. Phys.*, vol. 34, pp. 326–376, 1962.

[129] R. F. W. Bader and M. E. Stephens, "Spatial localization of the electronic pair and number distributions in molecules ," *J. Am. Chem. Soc.*, vol. 97, pp. 7391–7399, 1975.

[130] D. R. Hartree, "The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, no. 1, pp. 89–110, 1928.

[131] D. R. Hartree, "The calculation of atomic structures," *Rep. Prog. Phys.*, vol. 11, no. 5, pp. 113–143, 1947.

[132] E. H. Lieb, "Variational principle for many-fermion systems," *Phys. Rev. Lett.*, vol. 46, no. 7, pp. 457–459, 1981.

[133] T. Gilbert, "Hohenberg-Kohn theorem for nonlocal external potentials," *Phys. Rev. B*, vol. 12, no. 6, p. 2111, 1975.

[134] M. Piris and J. M. Ugalde, "Perspective on natural orbital functional theory," *Int. J. Quantum Chem.*, vol. 114, pp. 1169–1175, 2014.

[135] K. Pernal and K. J. H. Giesbertz, "Reduced Density Matrix Functional Theory (RDMFT) and Linear Response Time-Dependent RDMFT (TD-RDMFT)," in *Density-Functional Methods for Excited States* (N. Ferré, M. Filatov, and M. Huix-Rotllant, eds.), pp. 125–183, Springer International Publishing, 2016.

[136] M. Rodríguez-Mayorga, E. Ramos-Cordoba, M. Via-Nadal, M. Piris, and E. Matito, "Comprehensive benchmarking of density matrix functional approximations," *Phys. Chem. Chem. Phys.*, vol. 19, pp. 24029–24041, 2017.

[137] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, pp. B864–B871, 1964.

[138] M. Piris, "Natural Orbital Functional Theory," in *Reduced-Density-Matrix Mechanics: With Application to Many-Electron Atoms and Molecules* (D. A. Mazziotti, ed.), vol. 134 of *Advances in Chemical Physics*, ch. 14, pp. 385–427, Wiley Online Library, 2007.

[139] A. J. Coleman, "Structure of fermion density matrices," *Rev. Mod. Phys.*, vol. 35, no. 3, pp. 668–687, 1963.

[140] C. Garrod and J. K. Percus, "Reduction of the N-particle variational problem," *J. Math. Phys.*, vol. 5, no. 12, pp. 1756–1776, 1964.

[141] F. Weinhold and E. B. Wilson Jr, "Reduced density matrices of atoms and molecules. II. on the N-representability problem," *J. Chem. Phys.*, vol. 47, no. 7, pp. 2298–2311, 1967.

[142] D. A. Mazziotti, "Structure of fermionic density matrices: complete N-representability conditions," *Phys. Rev. Lett.*, vol. 108, no. 26, p. 263002, 2012.

[143] D. A. Mazziotti and R. M. Erdahl, "Uncertainty relations and reduced density matrices: Mapping many-body quantum mechanics onto four particles," *Phys. Rev. A*, vol. 63, no. 4, p. 042113, 2001.

[144] G. Gidofalvi and D. A. Mazziotti, "Molecular properties from variational reduced-density-matrix theory with three-particle N-representability conditions," *J. Chem. Phys.*, vol. 126, no. 2, p. 024105, 2007.

[145] R. M. Erdahl, "Representability," *Int. J. Quantum Chem.*, vol. 13, no. 6, pp. 697–718, 1978.

[146] J. M. Herbert and J. E. Harriman, "N-representability and variational stability in natural orbital functional theory," *J. Chem. Phys.*, vol. 118, no. 24, pp. 10835–10846, 2003.

[147] M. Piris, "A generalized self-consistent-field procedure in the improved BCS theory," *J. Math. Chem.*, vol. 25, no. 1, pp. 47–54, 1999.

[148] M. Piris, "Interacting pairs in natural orbital functional theory," *J. Chem. Phys.*, vol. 141, no. 4, p. 044107, 2014.

[149] J. W. Hollett, H. Hosseini, and C. Menzies, "A cumulant functional for static and dynamic correlation," *J. Chem. Phys.*, vol. 145, no. 8, p. 084106, 2016.

[150] J. W. Hollett and N. Pegoretti, "On-top density functionals for the short-range dynamic correlation between electrons of opposite and parallel spin," *J. Chem. Phys.*, vol. 148, no. 16, p. 164111, 2018.

[151] M. A. Buijse, *Thesis: Electron Correlation. Fermi and Coulomb holes, dynamical and nondynamical correlation.* PhD thesis, Vrije Universiteit, Amsterdam, The Netherlands, 1991.

[152] M. A. Buijse and E. J. Baerends, "An approximate exchange-correlation hole density as a functional of the natural orbitals," *Mol. Phys.*, vol. 100, pp. 401–421, 2002.

[153] G. Csányi and T. A. Arias, "Tensor product expansions for correlation in quantum many-body systems," *Phys. Rev. B*, vol. 61, no. 11, p. 7348, 2000.

[154] A. M. K. Müller, "Explicit approximate expression between reduced two- and one-particle density matrices," *Phys. Lett.*, vol. 105A, pp. 446–452, 1984.

[155] O. Gritsenko, K. Pernal, and E. J. Baerends, "An improved density matrix functional by physically motivated repulsive corrections," *J. Chem. Phys.*, vol. 122, no. 20, p. 204102, 2005.

[156] S. Goedecker and C. J. Umrigar, "Natural orbital functional for the many-electron problem," *Phys. Rev. Lett.*, vol. 81, pp. 866–869, 1998.

[157] V. N. Staroverov and G. E. Scuseria, "Optimization of density matrix functionals by the Hartree–Fock–Bogoliubov method," *J. Chem. Phys.*, vol. 117, no. 24, pp. 11107–11112, 2002.

[158] G. Csányi, S. Goedecker, and T. A. Arias, "Improved tensor-product expansions for the two-particle density matrix," *Phys. Rev. A*, vol. 65, no. 3, p. 032510, 2002.

[159] M. A. L. Marques and N. N. Lathiotakis, "Empirical functionals for reduced-density-matrix-functional theory," *Phys. Rev. A*, vol. 77, no. 3, p. 032509, 2008.

[160] M. Piris, "A new approach for the two-electron cumulant in natural orbital functional theory," *Int. J. Quantum Chem.*, vol. 106, pp. 1093–1104, 2006.

[161] M. Piris, X. Lopez, and J. M. Ugalde, "Dispersion interactions within the Piris natural orbital functional theory: The helium dimer," *J. Chem. Phys.*, vol. 126, no. 21, p. 214103, 2007.

[162] M. Piris, J. M. Matxain, X. Lopez, and J. M. Ugalde, "Communications: Accurate description of atoms and molecules by natural orbital functional theory," *J. Chem. Phys.*, vol. 132, p. 031103, 2010.

[163] M. Piris, J. M. Matxain, X. Lopez, and J. M. Ugalde, "Communication: The role of the positivity N-representability conditions in natural orbital functional theory," *J. Chem. Phys.*, vol. 133, p. 111101, 2010.

[164] M. Piris, X. Lopez, F. Ruipérez, J. M. Matxain, and J. M. Ugalde, "A natural orbital functional for multiconfigurational states.," *J. Chem. Phys.*, vol. 134, no. 16, p. 164102, 2011.

[165] M. Piris, J. Matxain, and X. Lopez, "The intrapair electron correlation in natural orbital functional theory," *J. Chem. Phys.*, vol. 139, no. 23, pp. 234109–234109, 2013.

[166] M. Piris and N. H. March, "Low-lying isomers of free-space halogen clusters with tetrahedral and octahedral symmetry in relation to stable molecules such as SF_6," *J. Phys. Chem. A*, vol. 119, no. 40, pp. 10190–10194, 2015.

[167] M. Piris, "Natural orbital functional for multiplets," *Phys. Rev. A*, vol. 100, no. 3, p. 032508, 2019.

[168] R. F. W. Bader and M. E. Stephens, "Fluctuation and correlation of electrons in molecular systems," *Chem. Phys. Lett.*, vol. 26, pp. 445–449, 1974.

[169] X. Fradera, M. A. Austen, and R. F. W. Bader, "The Lewis Model and beyond," *J. Phys. Chem. A*, vol. 103, pp. 304–314, 1999.

[170] E. Sim, S. Song, and K. Burke, "Quantifying density errors in DFT," *J. Phys. Chem. Lett.*, vol. 9, no. 22, pp. 6385–6392, 2018.

[171] A. J. Thakkar, "Extracules, intracules, correlation holes, potentials, coefficients and all that," in *Density Matrices and Density Functionals* (R. M. Erdahl and V. H. Smith Jr., eds.), pp. 553–581, Springer, Doerdrecht, 1987.

[172] R. Boyd and J. Ugalde, *Computational Chemistry part A*. PhD thesis, ed S Fraga (Amsterdam: Elsevier) p 273–299, 1992.

[173] A. S. Eddington, *Fundamental theory*. Cambridge: Cambridge University Press, 1946.

[174] P. Debye, "Scattering from non-crystalline substances," *Ann. Physik*, vol. 46, pp. 809–823, 1915.

[175] P. Debye, "Zerstreuung von Röntgenstrahlen und quantentheorie," *Physik Z.*, vol. 24, no. 8, pp. 161–166, 1923.

[176] P. Debye, "Note on the Scattering of X-rays," *J. Math. and Phys.*, vol. 4, no. 1-4, pp. 133–147, 1925.

[177] P. M. Gill, D. L. Crittenden, D. P. O'Neill, and N. A. Besley, "A family of intracules, a conjecture and the electron correlation problem," *Phys. Chem. Chem. Phys.*, vol. 8, no. 1, pp. 15–25, 2006.

[178] D. L. Crittenden and P. M. Gill, "Intracule Functional Theory," in *Solving The Schrödinger Equation: Has Everything Been Tried?*, pp. 1–23, World Scientific, 2011.

[179] T. Koga and H. Matsuyama, "Electron–electron coalescence and counterbalance densities for atoms in Hartree–Fock theory," *J. Chem. Phys.*, vol. 107, no. 23, pp. 10062–10066, 1997.

[180] T. Koga and H. Matsuyama, "Electron-electron coalescence and counterbalance densities of atoms in position and momentum spaces," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 30, no. 24, pp. 5631–5641, 1997.

[181] T. Koga, "Electron-electron counterbalance hole in Hartree–Fock theory," *J. Chem. Phys.*, vol. 108, no. 6, pp. 2515–2518, 1998.

[182] J. M. Mercero, J. E. Fowler, C. Sarasola, and J. M. Ugalde, "Atomic configuration-interaction electron-electron counterbalance densities," *Phys. Rev. A*, vol. 59, no. 6, pp. 4255–4258, 1999.

[183] E. Valderrama, J. M. Mercero, and J. M. Ugalde, "The separation of the dynamical and non-dynamical electron correlation effects," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 34, no. 3, p. 275, 2001.

[184] E. Valderrama, X. Fradera, and J. M. Ugalde, "Determination of the integrated X-ray scattering intensities through the electron-pair relative-motion density at the origin," *Phys. Rev. A*, vol. 64, no. 4, p. 044501, 2001.

[185] R. J. Weiss, *X-ray determination of electron distributions*, vol. 6. Amsterdam: North Holland Publishing Co., 1966.

[186] R. A. Bonham and M. Fink, *High Energy Electron Scattering*. New York: Van Nostrand Reinhold, 1974.

[187] A. J. Thakkar and V. H. Smith Jr, "Form factors and total scattering intensities for the helium-like ions from explicitly correlated wavefunctions," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 11, no. 22, p. 3803, 1978.

[188] A. J. Thakkar, A. Tripathi, and V. H. Smith, "Anisotropic electronic intracule densities for diatomics," *Int. J. Quantum Chem.*, vol. 26, no. 2, pp. 157–166, 1984.

[189] A. J. Thakkar, A. Tripathi, and V. H. Smith Jr, "Molecular X-ray and electron-scattering intensities," *Phys. Rev. A*, vol. 29, no. 3, p. 1108, 1984.

[190] N. A. Besley, A. M. Lee, and P. M. W. Gill, "Computation and analysis of molecular Hartree–Fock momentum intracules," *Mol. Phys.*, vol. 100, no. 11, pp. 1763–1770, 2002.

[191] N. A. Besley and P. W. Gill, "Atomic and molecular intracules for excited states," *J. Chem. Phys.*, vol. 120, pp. 7290–7297 (and references therein), 2004.

[192] X. Fradera, M. Duran, and J. Mestres, "The mapping of the local contributions of Fermi and Coulomb correlation into intracule and extracule density distributions," *J. Chem. Phys.*, vol. 113, no. 7, pp. 2530–2543, 2000.

[193] T. Koga and H. Matsuyama, "Explicitly correlated extracule densities for two-electron atoms," *Int. J. Quantum Chem.*, vol. 74, no. 5, pp. 455–465, 1999.

[194] R. J. Boyd, C. Sarasola, and J. M. Ugalde, "Intracule densities and electron correlation in the hydrogen molecule," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 21, no. 14, p. 2555, 1988.

[195] C. Sarasola, L. Dominguez, M. Aguado, and J. Ugalde, "The Laplacian of the intracule and extracule densities and their relationship to the shell structure of atoms," *J. Chem. Phys.*, vol. 96, no. 9, pp. 6778–6783, 1992.

[196] J. Cioslowski and M. Martinov, "Electronegativity equalization in polyyne carbon chains," *J. Phys. Chem.*, vol. 100, pp. 6156–6160, 1996.

[197] X. Fradera, M. Duran, and J. Mestres, "The relevance of the Laplacian of intracule and extracule density distributions for analyzing electron–electron interactions in molecules," *J. Chem. Phys.*, vol. 107, no. 9, pp. 3576–3583, 1997.

[198] J. Cioslowski and G. Liu, "Electron intracule densities and coulomb holes from energy-derivative two-electron reduced density matrices," *J. Chem. Phys.*, vol. 109, no. 19, pp. 8225–8231, 1998.

[199] J. Cioslowski and G. Liu, "Topology of electron-electron interactions in atoms and molecules. II. The correlation cage," *J. Chem. Phys.*, vol. 110, no. 4, pp. 1882–1887, 1999.

[200] M. Piris, X. Lopez, and J. Ugalde, "Correlation holes for the helium dimer," *J. Chem. Phys.*, vol. 128, no. 13, p. 134102, 2008.

[201] M. C. Per, S. P. Russo, and I. K. Snook, "Anisotropic intracule densities and electron correlation in H_2: A quantum Monte Carlo study," *J. Chem. Phys.*, vol. 130, no. 13, p. 134103, 2009.

[202] J. K. Pearson, P. M. Gill, J. M. Ugalde, and R. J. Boyd, "Can correlation bring electrons closer together?," *Mol. Phys.*, vol. 107, no. 8-12, pp. 1089–1093, 2009.

[203] E. Valderrama, J. Ugalde, and R. Boyd, *Many-Electron Densities and Reduced Density Matrices*. Kluwer Academic/Plenum Publishers New York, 2000.

[204] C. A. Coulson and A. H. Neilson, "Electron correlation in the ground state of helium," *Proc. Phys. Soc. London*, vol. 78, no. 5, pp. 831–837, 1961.

[205] R. F. Boyd and C. A. Coulson, "Coulomb hole in some excited states of helium," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 6, pp. 782–793, 1973.

[206] R. Boyd and C. Coulson, "The Fermi hole in atoms," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 7, no. 14, p. 1805, 1974.

[207] T. J. Lee, J. E. Rice, G. E. Scuseria, and H. F. Schaefer, "A diagnostic for determining the quality of single-reference electron correlation methods," *Theor. Chim. Acta (Berlin)*, vol. 75, p. 81, 1989.

[208] J. Cioslowski, "Density-driven self-consistent-field method: Density-constrained correlation energies in the helium series," *Phys. Rev. A*, vol. 43, no. 3, pp. 1223–1228, 1991.

[209] E. Valderrama, E. V. Ludeña, and J. Hinze, "Analysis of dynamical and nondynamical components of electron correlation energy by means of local-scaling density-functional theory," *J. Chem. Phys.*, vol. 106, pp. 9227–9235, 1997.

[210] E. Valderrama, E. V. Ludeña, and J. Hinze, "Assessment of dynamical and nondynamical correlation energy components for the beryllium-atom isoelectronic sequence," *J. Chem. Phys.*, vol. 110, no. 5, pp. 2343–2353, 1999.

[211] A. D. Gottlieb and N. J. Mauser, "New measure of electron correlation," *Phys. Rev. Lett.*, vol. 95, no. 12, p. 123003, 2005.

[212] U. R. Fogueri, S. Kozuch, A. Karton, and J. M. L. Martin, "A simple DFT-based diagnostic for nondynamical correlation," *Theor. Chem. Acc.*, vol. 132, no. 1, pp. 1–9, 2013.

[213] E. Proynov, F. Liu, and J. Kong, "Analyzing effects of strong electron correlation within Kohn-Sham density-functional theory," *Phys. Rev. A*, vol. 88, no. 3, p. 032510, 2013.

[214] A. Raeber and D. A. Mazziotti, "Large eigenvalue of the cumulant part of the two-electron reduced density matrix as a measure of off-diagonal long-range order," *Phys. Rev. A*, vol. 92, no. 5, p. 052502, 2015.

[215] C. L. Benavides-Riveros, N. N. Lathiotakis, C. Schilling, and M. A. Marques, "Relating correlation measures: The importance of the energy gap," *Phys. Rev. A*, vol. 95, no. 3, p. 032507, 2017.

[216] C. L. Benavides-Riveros, N. N. Lathiotakis, and M. A. Marques, "Towards a formal definition of static and dynamic electronic correlations," *Phys. Chem. Chem. Phys.*, vol. 19, pp. 12655–12664, 2017.

[217] S. Vuckovic, T. J. P. Irons, L. O. Wagner, A. M. Teale, and P. Gori-Giorgi, "Interpolated energy densities, correlation indicators and lower bounds from approximations to the strong coupling limit of DFT," *Phys. Chem. Chem. Phys.*, vol. 19, no. 8, pp. 6169–6183, 2017.

[218] A. Görling and M. Ernzerhof, "Energy differences between Kohn-Sham and Hartree-Fock wave functions yielding the same electron density," *Phys. Rev. A*, vol. 51, no. 6, p. 4501, 1995.

[219] E. Ramos-Cordoba, P. Salvador, and E. Matito, "Separation of dynamic and nondynamic correlation," *Phys. Chem. Chem. Phys.*, vol. 18, pp. 24015–24023, 2016.

[220] V. H. Smith Jr, "Approximate natural orbitals for carbon ^{1}S," *Theor. Chim. Acta (Berlin)*, vol. 7, p. 245, 1967.

[221] E. Ramos-Cordoba and E. Matito, "Local descriptors of dynamic and nondynamic correlation," *J. Chem. Theory Comput.*, vol. 13, pp. 2705–2711, 2017.

[222] M. Via-Nadal, M. Rodríguez-Mayorga, and E. Matito, "A salient signature of van der Waals interactions," *Phys. Rev. A*, vol. 96, no. 5, p. 050501, 2017.

[223] D. Carrascal, J. Ferrer, J. C. Smith, and K. Burke, "The Hubbard dimer: A density functional case study of a many-body problem," *J. Phys.: Condens. Matter*, vol. 27, no. 39, p. 393001, 2015.

[224] K. Deur, L. Mazouin, and E. Fromager, "Exact ensemble density functional theory for excited states in a model system: Investigating the weight dependence of the correlation energy," *Phys. Rev. B*, vol. 95, no. 3, p. 035120, 2017.

[225] N. R. Kestner and O. Sinanoglu, "Study of electron correlation in helium-like systems using an exactly soluble model," *Phys. Rev.*, vol. 128, p. 2687, 1962.

[226] E. Santos, "Cálculo aproximado de la energía de correlación entre dos electrons," *An. R. Soc. Esp. Fís. Quím.*, vol. 64, p. 117, 1968.

[227] M. Taut, "Two electrons in an external oscillator potential: Particular analytic solutions of a coulomb correlation problem," *Phys. Rev. A*, vol. 48, p. 3561, 1993.

[228] N. H. March and R. Pucci, "Asymptotic form of first-order density matrix for atoms and molecules," *J. Chem. Phys.*, vol. 75, no. 1, pp. 496–497, 1981.

[229] M. Ernzerhof, K. Burke, and J. P. Perdew, "Long-range asymptotic behavior of ground-state wave functions, one-matrices, and pair densities," *J. Chem. Phys.*, vol. 105, no. 7, pp. 2798–2803, 1996.

[230] A. Unsöld, "Quantentheorie des Wasserstoffmolekülions und der Born-Landéschen Abstoßungskräfte," *Z. Physik*, vol. 43, no. 8, pp. 563–574, 1927.

[231] T. Janowski and P. Pulay, "A benchmark comparison of σ/σ and π/π dispersion: the dimers of naphthalene and decalin, and coronene and perhydrocoronene," *J. Am. Chem. Soc.*, vol. 134, no. 42, pp. 17520–17525, 2012.

[232] S. Goedecker and C. J. Umrigar, "Natural orbital functional theory," in *Many-Electron Densities and Reduced Density Matrices*, ch. 8, pp. 165–181, New York: Springer, 2000.

[233] E. Ramos-Cordoba, X. Lopez, M. Piris, and E. Matito, "H_4: A challenging system for natural orbital functional approximations," *J. Chem. Phys.*, vol. 143, no. 16, p. 164112, 2015.

[234] M. Piris, X. Lopez, and J. M. Ugalde, "Electron-pair density relaxation holes," *J. Chem. Phys.*, vol. 128, no. 21, p. 214105, 2008.