DOCTORATE PROGRAM IN THEORETICAL CHEMISTRY
AND COMPUTATIONAL MODELLING

APRIL, 2017

FROM BONDING TO MOLECULAR PROPERTIES IN THE CONTEXT OF QUANTUM CHEMICAL TOPOLOGY

DANIEL MENÉNDEZ CRESPO

UNIVERSITY OF OVIEDO.
Physical and Analitical Chemistry Department.
Physical Chemistry subject area.
Quantum Chemistry Group.

PROGRAMA DE DOCTORADO EN QUÍMICA TEÓRICA Y MODELIZACIÓN COMPUTACIONAL

ABRIL, 2017

DESDE EL ENLACE A LAS PROPIEDADES MOLECULARES EN EL CONTEXTO DE LA TOPOLOGÍA QUÍMICO CUÁNTICA

DANIEL MENÉNDEZ CRESPO

UNIVERSIDAD DE OVIEDO.
Departamento de Química Física y Analítica.
Área de Química Física.
Grupo de Química Cuántica.

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español/Otro Idioma: Desde el enlace a las propiedades moleculares en el contexto de la Topología Químico Cuántica.	Inglés: From bonding to molecular properties in the context of Quantum Chemical Topology.

2.- Autor	
Nombre: Daniel Menéndez Crespo	DNI/Pasaporte/NIE:
Programa de Doctorado: Química Teórica y Modelización Computacional	
Órgano responsable: Centro Internacional de Postgrado	

RESUMEN (en español)

El enlace químico podría ser considerado como el pilar central de la química. Sin embargo, a razón de no ser un observable mecánico cuántico, su comprensión escapa a un fundamento teórico riguroso. En este escenario, los acontecimientos históricos, y no el razonamiento científico, han condicionado los modelos prevalecientes sobre el enlace químico. Es de esta manera que la teoría de orbitales moleculares (MO) ha alcanzado su estado actual, convirtiéndose en una teoría tan arraigado en la química moderna que gran parte del vocabulario químico proviene de ella. Sin embargo, la teoría MO se basa extrañamente en objetos que viven en un espacio multidimensional complejo que rara vez evoca la intuición química natural, que consiste en considerar a los electrones como entidades que viven en el espacio real. Bajo esta premisa, ha florecido una teoría del enlace químico en el espacio real que tiene como un claro exponente la Teoría Cuántica de los Átomos en las Moléculas (QTAIM) propuesta por Richard Bader y colaboradores. Basado en la partición espacial propuesta por QTAIM, destacamos en esta tesis la posibilidad de realizar una descomposición energética (IQA), medir las probabilidades de las posibles poblaciones de electrones en la cada región QTAIM (EDFs), y de explorar imágenes efectivas de un electrón válidas para sistemas correlacionados que imitan a los del paradigma MO (NAdOs).

Específicamente, hemos enfatizado la idoneidad de la partición de energía IQA para definir una energía de enlace teóricamente sólida, llamada energía de enlace in situ. Esto combinado con las otras herramientas mencionadas anteriormente nos permite conocer, en circunstancias especiales, cuál es el estado de valencia de los fragmentos moleculares y cómo se forman los componentes de unión, todo esto contribuyendo a una comprensión muy íntima de su comportamiento electrónico en equilibrio así como en las diferentes etapas de la formación de los enlaces. Teniendo en cuenta las tendencias hacia sistemas cada vez más grandes y el escalado particularmente caro de IQA hasta ahora, en parte debido al cálculo de la correlación de intercambio entre dos cuencas diferentes, también hemos propuesto llevar a cabo una aproximación multipolar de este término, del mismo modo que para la energía de interacción de Coulomb. Se ha demostrado que este enfoque es preciso, incluso con un truncamiento que considera como máximo términos de interacción carga-cuadrupolo, si las cuencas que interactúan están lo suficientemente alejadas unas de otras. El enfoque ha sido probado con una variada selección de moléculas. También se establece la conexión del primer término de la expansión y uno de los descriptores más importantes en QTAIM, el índice de deslocalización.

También se puede demostrar que los descriptores de enlace alojados en el marco de la Topología Química Cuántica (QCT) están implicados en la explicación de una clase más amplia de fenómenos químicos. En esta tesis se ha señalado que existe un vínculo entre los índices de enlace y el tensor de localización utilizado en la teoría moderna del estado aislante. El último señala un comportamiento aislante o conductivo basándose en sus propiedades de convergencia o divergencia en el límite termodinámico. Después de una partición del espacio, hemos demostrado que la convergencia/divergencia del tensor sólo depende de los componentes interatómicos que a su vez están dominados por el índice de deslocalización. De este modo, se obtiene una noción químicamente atractiva del tensor de localización. Otro tema de interés que hemos tratado es el estudio de las interacciones débiles en los sólidos moleculares. Para ello, hemos aprovechado las propiedades topológicas que presenta el potencial electrostático. Debido a que el trabajo previo en la literatura sobre la topología del potencial electrostático en sólidos es escaso, hemos emprendido primero una exploración de sus características en el complejo de carga BTDMTTF-TCNQ. Las interacciones predichas tanto por la densidad y el potencial electrostático se buscaron exhaustivamente, siendo más tarde entrelazadas para proporcionar una mejor comprensión del empaquetamiento del cristal. También, a partir de la partición combinada del espacio, hemos descifrado cuáles son los actores principales que conducen la transferencia de la carga.

RESUMEN (en Inglés)

The chemical bond might be considered as the central pillar of chemistry. Not being a quantum mechanical observable, however, its understanding escapes a rigorous theoretical foundation. In this scenario, historic events, not scientific reasoning, have conditioned the prevailing models on chemical bonding. It is in this way that molecular orbital (MO) theory has achieved its present status, becoming so rooted in modern chemistry that much of the chemist vocabulary comes from it. Taken from a non-MO biased perspective, however, MO theory is strangely based on objects that live on a complex multidimensional space that rarely evokes the natural chemical intuition, made up of considering electrons as entities living in real space. Under this premise a theory of chemical bonding in real space has flourished that has as a clear exponent in the Quantum Theory of Atoms in Molecules (QTAIM) proposed by Richard Bader and coworkers. Based on the space partition proposed by QTAIM, we highlight in this Ph.D. thesis the possibility of performing an energy partitioning (IQA), of measuring the probabilities of the possible electron populations in the QTAIM regions (EDFs), and of exploring effective one-electron images valid for correlated systems that mimic those of the MO paradigm (NAdOs).

Specifically, we have emphasized the suitability of the IQA energy partition to define a theoretically sound bond energy, called in situ bond energy. This combined with the other tools mentioned above allows us to know, under special circumstances, what the valence state of the molecular fragments are and how the binding components are formed, all this contributing to a very intimate understanding of both their electronic behavior in equilibrium as well as at the different stages of the formation of bonds. Considering the trends towards ever larger systems and the particularly expensive scaling of IQA so far, in part due to the calculation of the exchange-correlation between two different basins, we have also proposed to carry out a multipolar approximation of this term, in the same fashion as for the Coulombic interaction energy. This approach has been shown to be accurate, even with a truncation to at most charge-quadrupole interaction terms, if the interacting basins are far enough from each other. The approach has been tested with a varied selection of molecules. Also, the connection of the first term of the expansion and one of the most important descriptors in the QTAIM, the delocalization index, is also established.

Bonding descriptors housed under the framework of Quantum Chemical Topology (QCT) can also be shown to be involved in the explanation of a broader class of chemical phenomena. In this thesis, it has been pointed out that a link exists between bond order indices and the localization tensor used in the modern theory of the insulating state. The last one signals insulating or conducting behavior based on its convergence or divergence properties in the thermodynamical limit. After a partitioning of space we have demonstrated that convergence/divergence of the tensor depends only on interatomic components that in turn are dominated by the delocalization index. Thus a chemically appealing notion of the localization tensor is gained in the process. Another topic of interest that we have deald with is the study of weak interactions in molecular solids. For this, we have taken advantage of the topological properties that the electrostatic potential presents. Because the previous work in the literature on the topology of the electrostatic potential in solids is scarce, we have undertaken first an exploration of its characteristincs in the charge-complex BTDMTTF-TCNQ. The interactions predicted both by the density and the electrostatic potential were searched exhaustively, being later intertwined to provide a better understanding of the crystal packaging. Also from the combined partition of space, we deciphered which are the main actors driving the charge transfer.

AGRADECIMIENTOS

A lo largo del camino transcurrido para la defensa de esta tesis he conocido a mucha gente maravillosa, con la que además he compartido muy buenos momentos. Primero y ante todo, me gustaría agradecer sinceramente a mis directores de tesis — Aurora y Ángel — su esfuerzo, dedicación, e insistencia para transmitirme sus conocimientos. Gracias por estar ahí para todo.

En especial, a Víctor le agradezco que haya infundido en mí la curiosidad y el optimismo necesario para seguir adelante. Así como de Evelio admiro su mano diestra y resolución en los momentos en los que he reclamado su ayuda. Igualmente, a Michi le agradezco haberme introducido en el conocimiento de la cuántica cuando era una gran desconocida para mí. En definitiva, a los profesores del grupo de Oviedo les debo mi formación como investigador y, en cierta medida, como persona.

Con especial cariño me acuerdo de aquellos con los que más tiempo he pasado, me refiero a mis compañeros de laboratorio. Empezando por Alfonso, a quien da gusto verle entrar por la puerta con una sonrisa en la cara y diciendo: "¡Eyyy, chavales!". A todos nos encanta que nos contamines de alegría. De José Luis, que ahora se encontrará en México pasándoselo pipa, seguro. Fernando Izquierdo y José Manuel, a los que les doy ánimo porque ya os queda poco. Por cierto, a Fernando y Serra os deseo una feliz boda. Nico, aunque estuviste por poco tiempo dejaste una impronta que no olvidaremos. También me acuerdo de Hussien. Y los integrantes más noveles, Carlos y Fernando Grávalos, a quienes animo porque con toda seguridad quedarán enganchados. Así como aquellos que en estos momentos os encotrais en el extranjero: David, Marcos, Roberto, Mamel, Miriam, Alberto, Julia, Marian, pero que siempre aprovecháis para realizar una visita. Gracias por crear entre todos un ambiente de diálogo sobre temas que incluso trascienden los quehaceres de la ciencia. Siempre es interesante escuchar vuestras opiniones en las tertulias matutinas. ¡A ver cuando organizamos de nuevo una cena o comida!

A los compañeros de máster solo les digo que espero verles en el futuro.

A todos les deseo lo mejor.

ACRONYMS

AIM	Atoms In Molecules theory
BCP	Bond Critical Point or $(3,-1)$ ρ-**CP**
BTDMTTF	bis(thiodimethylene)–tetrathiafulvalene
CCD	Charge-Coupled Device
CP	Critic Point
CUBE	`Gaussian` parallepiped grid 3D data file
`critic2`	A program for the topological anaylisis of real-space scalar fields in periodic systems
CWM	Constrained Wavefunction Method
DFT	Density Functional Theory
EBCP	Electrostatic Bond Critical Point or $(3,-1)$ ϕ-**CP**
ED	Electronic Density (one body electron density)
ELF	Electron Localization function
ELI	Electron Localizability Indicator
EOS	Equation Of State
ESP	Electrostatic Potential (one body electrostatic potential)
`espresso`	opEn Source Package for Research in Electronic Structure, Simulation, and Optimization
`Gaussian`	Electronic structure code
HF	Hartree Fock Optimization of a multielectronic system monoelectronic orbitals through minimization of a Russell-Saunders state energy under spin-orbitals orthonormalization condition.
HK	Hohenberg–Kohn theorems
IAM	Independent Atom Model
IAS	Inter Attractor Surface

IUCR International Union of Crystallography

KS Kohn–Sham equations

LSEXP Multipolar refinement program by Hirshfeld

LATITUDINAL Latitudinal pseudo-isolated dimer

LHS left–hand side

LONGITUDINAL Longitudinal pseudo-isolated dimer

MEM Maximum Entropy Method

MEP Molecular Electrostatic Potential

MOLLY Multipolar refinement program by Hansen and Coppens

MoPro Cristallographic least square refinement package allowing structural or charge density studies of cristal stuctures of variable sizes, ranging from small molecules to biological macromolecules.

MoProViewer A graphical interface to MoPro

NCI Non Covalent Interactions

Orca Electronic structure code

POP Multipolar refinement program by Craven, Weber, He

QE Quantum espresso equations

Plane waves + pseudopotentials. Electronic structure code for extended systems.

QCT Quantum Chemical Topology

QTAIM Quantum Theory of Atoms in Molecules

MPn Moller Plesset order n

RDG Reduced Density Gradient

RHS right–hand side

RDM Reduced Density Matrix

TTF tetrathiafulvalene

TCNQ 7,7,8,8–tetracyanoquinodimethane

UNESCO United Nations Educational, Scientific and Cultural Organization

UPF Unified Pseudopotential File

Pseudopotential standard file format.

VSCC	Valence Shell Charge Concentration
VSCD	Valence Shell Charge Depletion
VSEPR	Valence Shell Electron Pairs Repulsion
VMoPro	Scalar fields analysis tool
XDM	Exchange Dipole Moment
XRD	X–Ray Diffraction

CONTENTS

INTRODUCTION

The nature of chemical bonding remains one the most important unanswered questions in chemistry, joining an interdisciplinary community that tries to disentangle the reasons leading to bonding formation/breaking and their prediction to build new materials, understand biochemical processes, and control material properties. Much of the technological advances are guided by new and improved knowledge of the nanostructure of matter and its implication at the macroscopic level.

The ability to analyze experimental and theoretical densities with a quantum mechanically grounded theory has led to a great progress of charge density analysis during the last decades. Books [66, 117, 179, 180] and references therein show an overview of the last decade. The density is a fundamental variable of nature as stated through the first Hohenberg–Kohn (HK) theorem. It is the *lingua franca* of X–Ray Diffraction (XRD), Density Functional Theory (DFT) and Quantum Theory of Atoms in Molecules (QTAIM).

However, the situation is far from desired. The information interchanged is limited to the one-particle density. Though the topological partition of the electronic density provides descriptors soon identifiable from chemical intuition, their meaning isn't still fully known and the information retrieved isn't complete enough to fully describe the bonding nature. For example, bond orders are highly dependent on the correlation. Therefore new descriptors not coming from the density but from the Reduced Density Matrices (n–RDM) are used instead to complement the results of QTAIM. The understanding of their nature becomes a task of prime importance.

The structure of this thesis is the following: In the first chapters a throughout introduction of the main theories of the computational modelling of molecules and materials is presented. Then, in chapter THEORY OF CHEMICAL BONDING we deal with the theory of chemical bonding in real space. The results are shown in part II and part III. On one hand, in part II we first treat the theme of covalent interactions, trying to asses the, often intrincate, appropriate quantification of bonding strength in chapter TOWARDS A UNIQUE MEASURE OF BOND STRENGTH THROUGH INTRINSIC BOND ENERGIES and second adopting a multipolar approximation in the realm of IQA energetic partitioning, chapter MULTIPOLAR EXPANSION OF THE EXCHANGE-CORRELATION INTERACTION ENERGY.

On the other hand, in part III we extend the application of current tools of Quantum Chemical Topology as well as broadening the set of tools. In chapter A CHEMICAL VIEW OF THE LOCALIZATION TENSOR we build a bridge between bond order descriptors and the locazation tensor. Finally, the chapter titled TOPOLOGY OF THE ELECTROSTATIC POTENTIAL IN SOLIDS translates the study of the electrostatic potential to extended phases with the aim of under-

standing its topological characteristics. There are many works published about the electrostatic potential but only a few analyze the topology in molecules, and even fewer in solids [113, 116, 187]. We considered appropriate the use of the electrostatic potential to study an organic conductor in crystal phase with the aim of extending the previous work by Mata et al. Also, among the many Quantum Chemical Topology (QCT) descriptors already present, aside from the well studied density, the electrostatic potential is the easier target to look at as a new source of chemical information, specially interesting are its applications in weakly interacting systems to approximate and predict long range interactions.

Part I

Methodology

1 COMPUTATIONAL QUANTUM CHEMISTRY

CONTENTS

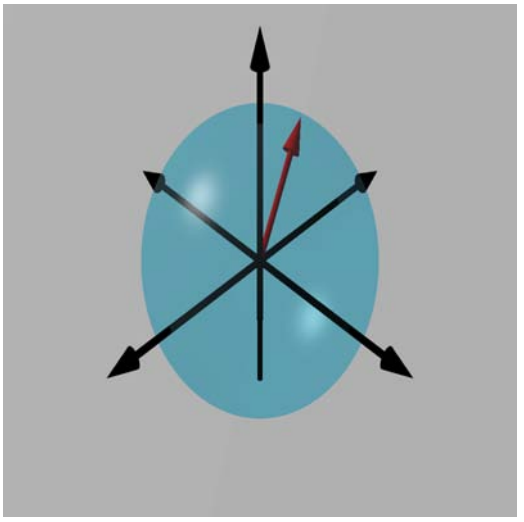


Illustration of a normalized pure state represented as a ray (red) in the unit Hilbert sphere. Otherwise, mixed states are the result of coupling pure states and are represented as vectors inside the sphere. In this neat description, evolution is represented by a path in the hypersphere.

All quantities reported will be in atomic units $\hbar = m_e = 4\pi\varepsilon_0 = e = 1$, otherwise the units used will be stated.

1.1 INTRODUCTION

The state of a N-particle quantum system is defined by a unit norm vector $|\Psi\rangle$ in the abstract Hilbert (unit) sphere, that is represented in the position space as

$$\Psi = \langle \{\boldsymbol{x}\}_{i=1}^{N} | \Psi \rangle. \tag{1–1}$$

The non-relativistic time evolution of this state is controlled by a unitary transformation conserving the number of particles of the system induced by H, the Hamiltonian:

$$i\frac{\partial \Psi}{\partial t} = \mathcal{H}\Psi. \tag{1–2}$$

From this master equation, named after Schrödinger, many properties of quantum systems can be discerned: (i) The superposition of any two solutions is again a solution. This shows that the set of solutions of this differential equation form a linear space. (ii) Any measurement should leave the system in another possible state, maybe the same, so the natural way to express a measurement is through a map that takes as an argument any possible state and returns another possible state, that is, through an operator. It turns out that measurements have associated an operator that is indeed linear and self-adjoint. (iii) In addition it is deduced from the Schrödinger equation that for particles under a Coulomb potential the wavefunction describing the system can only be complex valued. If it were real the RHS of equation (1–2) would be real because the Hamiltonian ($H = -\frac{\Delta^2}{2} + V(\mathbf{r}, t)$) has no imaginary terms, while the LHS has an explicit dependence on i.

For a stationary system (i.e. with a time independent Hamiltonian because there is no external time-varying field) the wavefunction at any time may be written as

$$\Psi(t) = \sum_i c_i e^{-iE_i t} \psi_i, \tag{1–3}$$

so any eigenfunction of the Hamiltonian can be expressed as the product of a unitary operator $U(t)$ and the initial wavefunction

$$\Psi(t) = e^{-iHt}\Psi(0) = U(t)\Psi(0). \tag{1–4}$$

With this factorization of time and spatial variables we can approximate the master equation as two independent equations leaving a time independent equation, equivalently written as an eigenvalue equation

$$-\sum_i^N \frac{\partial^2 \Psi}{2\partial \mathbf{r}_i^2} + V(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N)\Psi = E\Psi \qquad \text{or} \qquad H\Psi = E\Psi. \tag{1–5}$$

The eigenvalues and eigenfunctions of the Hamiltonian are the quantized energies and states the system can have. The state of the system at any time is known if we have all the eigenfunctions and eigenvalues of the Hamiltonian. Particularly important is the lowest eigenvalue. Guided by the minimum action principle a quantum chemical system will radiate energy to the surroundings until it reaches the *ground state*, i.e. the lower energy attainable.

If the Hamiltonian is real it suffices to contemplate real valued wavefunctions (from a superposition with the conjugate transpose $\Re(\Psi) = \frac{1}{2}(\Psi + \Psi^*)$) to find bounded states. Under a potential that contains only multiplicative factors, most of them are, the Hamiltonian contains a second order derivative as part of the kinetic energy operator. As the second derivative of the wavefunction is determined by Ψ and the third and successive derivatives only require the knowledge of the function and its first derivative, then the function itself could be known in the whole space from knowing exactly both entities at a single point and further propagating the solution through space. Another completely different thing is to perform this task, but the previous statement reflects how powerful this equation is. This observation has to do with the effort of some researchers of extracting chemical information from a subset of points in real space, i.e. critical points of the density or the source function.

Molecular systems are the pinnacle of quantum mechanics applications, requiring the simulation of hundreds or thousands of electrons simultaneously interacting among them and the nuclei all together. The computational cost of a direct attack of the problem increases exponentially, becoming rapidly outside of reach for current computational resources. Except for really simple systems, we must be content with approximate solutions. There are four main approximations at the time of modeling a molecular system:

- The system. We assume that the molecular system under study is isolated as if it were in the gas phase. In the next section we will reduce the molecular system to one with only electrons in a field of fixed nuclei. A further simplification is to model only valence electrons approximating the core electrons from previous calculations with what we call a pseudopotential.

- The model used to describe a system is encoded in the Hamiltonian. Particles are usually reduced to points in space without any structure. All known fundamental forces of nature involve two particle interactions, consequently a Hamiltonian of this type only contains one– and two–body operators. Molecular systems are mainly dominated by electromagnetic forces unless heavy elements are included. In such cases the electrons occupy high kinetic energy states with velocities that are a significant fraction of the speed of light, a (special) relativistic framework with scalar, two or four components is more adequate. For a non relativistic system

($c \to \infty$) of n electrons with clamped nuclei (infinite massive nuclei), the two-particle Coulomb Hamiltonian is:

$$H_{el} = \sum_{i=1}^{n} \left(-\frac{1}{2}\nabla_i^2 + v(\mathbf{r}_i) \right) + \frac{1}{2} \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{1–6}$$

- Another is the restriction of the space of possible solutions that is briefly described in sections 1.3, 1.4,and 1.6. We reduce the space from where we will extract our final wavefunction.

- Finally in section 1.5 we deal with the discretization of the high dimensional problem to a finite basis.

1.2 DECOUPLING OF ELECTRONIC AND NUCLEAR MOTION

In a system constituted by M nuclei with an atomic number Z_A each and N electrons, each particle has its own contribution to the total kinetic energy $T_e + T_n$ and interacts with its equals $V_{ee} + V_{nn}$ as well as with its unequals $V_{en} + V_{ne}$ through a Coulomb potential. Paying attention to double counting of equivalent interactions the non-relativistic molecular Hamiltonian (in position space) is

$$H_{mol} = T_e + V_{ne} + V_{ee} + T_n + V_{nn} \tag{1–7}$$

$$= \sum_{i}^{N} \left(-\frac{1}{2}\Delta_{\mathbf{r}_i} - \sum_{A}^{M} \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \frac{1}{2} \sum_{\substack{j \\ j \neq i}}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \tag{1–8}$$

$$+ \sum_{A}^{M} \left(-\frac{1}{2m_A}\Delta_{\mathbf{R}_A} + \frac{1}{2} \sum_{\substack{B \\ B \neq A}}^{M} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \right), \tag{1–9}$$

$$\tag{1–10}$$

the result of replacing in the classical Hamilton function the coordinates and the momentum by the linear operators $\mathbf{r}$ and $-i\nabla$. The Hamiltonian operator only has a finite spectrum if we get rid of the translational part after separating the center of mass and the internal motion

$$H_{mol} = H_{translation} + H_{internal}. \tag{1–11}$$

As the Hamiltonian does not operate on nuclear spin variables and nuclei are distinguishable the total wavefunction depends on electronic spin-spatial coordinates $\mathbf{x} = (\mathbf{r}, \sigma)$ and nuclear spatial coordinates $\mathbf{R}$ alone

$$\Psi(\underline{\mathbf{x}}, \underline{\mathbf{R}}). \tag{1–12}$$

The spatial variables can have any value in $\mathbb{R}^3$ and z-axis projected electronic spin can only take two eigenvalues: a positive or negative $\pm s_z$ that we codify with the domain $\{+, -\}$. Here we condense in $\underline{\mathbf{x}}$ the tuple of vectors $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ of all particles. The dimensionality of the problem is too high to be solved: it grows exponentially with exponent $6(N + M)$. Consequently, it is our duty to seek a way to reduce the high number of independent variables available using physical insight.

The motion of atoms in the course of a chemical reaction is in some aspects analogous to the motion of sailing boats in a regatta. Wind and sea waves turn and move often at speeds higher than the boat, becoming force directors of the boats motion. For us the boats are the nuclei and the sea and atmosphere are the electronic cloud. In the spirit of this imperfect analogy a ground state system close to stability is a collection of boats (nuclei) floating in a calm sea. Any turbulence created by the boats is instantaneously recognized by the system air/sea that automatically acts upon the changes introduced, whereas the boat feels the average force exerted by the particles that constitute water and air.

Similarly, events that take place inside matter are distributed in a wide range of time scales. Solvation, chemical reactions, rotations of chemical groups, vibrations, inter-system crossing, internal conversion, electronic relaxation, ... are all relevant processes that determine the properties of matter. The lighter the particles involved the shorter the time scale except when quantum selection rules apply. Thus, for events were electrons are involved shorter time scales apply. Taking into account that the fastest type motion a nucleus can achieve is a vibration, we will check that the timescales of nuclear vibrations are usually greater than for electronic reorganizations. A treatment of nuclei and electrons as oscillators reveals that their relative frequencies of motion would be proportional to their relative weights, electrons move $\sqrt{\frac{m_N}{m_e}} \approx 100$ times faster than nuclei. For a more quantitative analysis on the effect of separating both see [26, 27]. Therefore, one can examine the electronic distribution considering the nuclei as being fixed. This argument is the basis for the Born–Oppenheimer approximation.

Lets assume we can write the wavefunction as a product of a nuclear wavefunction and an electronic wavefunction

$$\Psi = \Psi_{el}\Psi_{nuc}. \tag{1–13}$$

To reduce the molecular problem to an electronic problem we have to take the asymptotic limit of infinitely heavy nuclei $m_A \to \infty$. Therefore the kinetic energy of nuclei vanishes. With an enough localized nuclear function at positions

$\underline{\mathbf{R}}$ the minimization problem is reduced to the minimization of the electronic energy

$$E(\underline{\mathbf{R}}) = \min_{\Psi_{el}} \left[\langle \Psi_{el} | H_{el} | \Psi_{el} \rangle \right] + \sum_{A}^{M} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \tag{1–14}$$

where the electronic wavefunction satisfies the normalization condition and

$$H_{el} = \sum_i \left(-\frac{1}{2}\Delta_i - \sum_A \frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} \right) + \sum_{i>j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} = \sum_i h_i + \sum_{i>j} g_{ij}. \tag{1–15}$$

Within this approximation, solving the molecular problem is equivalent to solving the electronic structure for each geometrical disposition of the nuclei. Semiclassical nuclei dynamics can be recovered letting the nuclear positions evolve on top of the potential energy (hyper)surface (PES) provided by (1–14) at various molecular conformations. Ψ_{el} depends now parametrically in the nuclei coordinates

$$\Psi_{el}(\underline{\mathbf{x}}; \underline{\mathbf{R}}). \tag{1–16}$$

The nuclear-nuclear repulsion energy is thus simply a constant that can be computed as a classical contribution introduced by point particles for a given nuclei geometry.

The potential exerted by nuclei on electrons and any external field is the so–called external potential.

Except where there are nearly degenerated states or we are close to crossings of electronic states the approximation works reasonably well. When a vibrational-electronic (vibronic) coupling is present the nuclear kinetic energy contains coupling terms that depend on derivatives of the electronic wavefunction.

From now on every time we mention Ψ it should be understood as Ψ_{el}.

1.3 MANY–ELECTRON WAVE FUNCTIONS

An electronic wavefunction $\Psi(\underline{\mathbf{x}}, \underline{\mathbf{R}})$ ought to describe a system of N electrons has to satisfy some restrictions to have physical meaning. Based on the physical interpretation proposed by M. Born that assures that the interpretation of this object comes through its complex square, $|\Psi(\{\mathbf{x}\})|^2$, that conveys the probability density for finding one electron at position $\mathbf{r}_1$ with spin σ_1, another at $\mathbf{r}_2$ with spin σ_2, and so on, the integral of this square should remain finite and normalized in order to represent a unit probability of finding all of the N particles of the system in the whole space $\|\Psi\| = 1$. We will assume that our wavefunctions are always normalized. The distribution has to be defined everywhere in space (continuous) and the kinetic energy has to be bounded to a finite value. Collecting these requirements it may be proven that the wavefunction dwells in a Hilbert space $H^1(\mathbb{R}^{3N} \times \{+,-\}^N)$.

A drastic simplification of the problem at hand is to factorize the wavefunction even further as a Hartree product of one-electron spin-orbitals, $\prod_{i=1}^{N} \psi_i(\mathbf{x}_i)$. Spinorbitals are the product of a spacial (orbital) and a spin function $\psi_i(\mathbf{x}) = \phi_i(\mathbf{r})s_i(\sigma)$ where each orbital can accommodate up to two electrons with different spin functions attached, $s_i = \alpha$ and $s_i = \beta$, that satisfy

$$\alpha(+) = \beta(-) = 1 \qquad \alpha(-) = \beta(+) = 0. \tag{1-17}$$

The way orbitals are constructed will be explained later in section 1.5. Notice that each electron has a well defined spin state. This factorization leads to a probability density that is simply the product of their squares

$$|\psi_1(\mathbf{r}_1)|^2|\psi_2(\mathbf{r}_2)|^2|\psi_3(\mathbf{r}_3)|^2\cdots$$

The statistical interpretation dictates that no correlation exists in the motion of electrons in this case. Conversely is not enough to describe a many-fermion system.

Furthermore, it must be borne in mind that when identical particles are described by quantum mechanic rules, they are indistinguishable. The essence relies in that the coordinates we are working with are not the actual coordinates but the independent variables of the probabilistic distribution function. This amounts to stating that any permutation of the spatial and spin coordinates of any set of electrons does not change the physical state (the probability density). Since the permutation operator is idempotent there are only [1] two representations of the wavefunction that obey this principle: (i) if the wavefunction is totally symmetric under this operation, we are talking of particles denominated bosons and, (ii) if the wavefunction is totally antisymmetric after applying the interchange

$$\Psi(\cdots,\mathbf{x}_i,\cdots,\mathbf{x}_j,\cdots) = -\Psi(\cdots,\mathbf{x}_j,\cdots,\mathbf{x}_i,\cdots) \tag{1-18}$$

we are dealing with fermions, particles with half integer spin such as electrons. Therefore the wavefunction can have negative amplitudes. It follows from this condition Pauli's exclusion principle, which asserts that two electrons with the same quantum numbers cannot be found in the same state.

Since fermionic wavefunctions must be antisymmetric under particle interchange , they can no longer be built as a direct product of independent one-electron functions, but instead as a tensor product keeping only antisymmetric

1 To be precise this argument is weak. Symmetric restrictions do not follow from indistinguishablity alone nor from relativistic invariance but we can adopt an additional symmetrization postulate [133] to close the loop.

terms. The simplest solution — as proposed by V. A. Fock — is to use Slater determinants, i.e. antisymmetrized Hartree products:

$$\Psi(\underline{\mathbf{x}}) = \mathcal{A}\prod_{i=1}^{N}\psi_i(\mathbf{x}_i) = \bigwedge_{i=1}^{N}\psi_i(\mathbf{x}_i) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_2) \\ \psi_1(\mathbf{x}_3) & \psi_2(\mathbf{x}_3) & \cdots & \psi_N(\mathbf{x}_3) \\ \psi_1(\mathbf{x}_4) & \psi_2(\mathbf{x}_4) & \cdots & \psi_N(\mathbf{x}_4) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \tag{1–19}$$

$$= \frac{1}{\sqrt{N!}}\det|\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\cdots\psi_N(\mathbf{x}_N)| = |\psi_1\psi_2\cdots\psi_N\rangle. \tag{1–20}$$

We thus define the wavefunction to belong to the set of Slater determinants [2] with all possible spin-orbitals under the condition of orthonormality:

$$\mathcal{S}_N = \left\{|\psi_1\psi_2\cdots\psi_N\rangle : \psi_i \in H^1(\mathbb{R}^3\times\{+,-\}), \langle\psi_i|\psi_j\rangle = \delta_{ij}\right\}. \tag{1–21}$$

A determinantal function is the least correlated [3] state of an electronic system, for it includes only some spatial *(Fermi) correlation* that forbids electrons of equal spin occupy the same positions. However, it provides a starting point to approximate the essential features of a many-electron problem. In contrast with the independent model (Hartree product) the correlation introduced reduces the mutual repulsion energy of parallel spin electrons, $\sum_{i>j}\langle\Psi|g_{ij}|\Psi\rangle$, but it also induces a nodal structure in the wavefunction and the kinetic energy therefore increases — it is a steeper function. The missing correlation arises from the Coulomb interactions and is named accordingly *Coulomb correlation*. For higher accuracy it is necessary to describe appropriately the simultaneous influence in the motion of each electron of all the others. This may be achieved by constructing linear combinations of Slater determinants constructed from a large set of one particle states. In fact, it may be shown that the correct wavefunction is reached with an infinite (convergent) sum of determinants constructed from a complete set of orbitals.

Antisymmetrized Hartree products, as a single determinant, are not always spin eigenfunctions. However, it is always possible to combine determinants constructed from the same spin-orbitals, to correct the problem introducing a new component of Fermi correlation that arises from spin symmetry. Electronic configurations free of this type of correlation, and thus single-determinant situations are: (i) closed shell systems, for each occupied $\psi_i(\mathbf{r},+)$ there is also an occupied $\psi_i(\mathbf{r},-)$ and/or (ii) high-spin systems where single occupied orbitals have electrons of the same spin. In situations where there are quasi-degenerate eigenstates, being two or more eigenstates energetically similar,

2 The columns of the determinant encode the list of orbitals and similarly the rows describe the electrons.

3 For clarification, correlation has two possible interpretations: (i) one is the statistical dependence of particles in the sense of Wigner and Steitz and (ii) other is the energetic difference with respect to a reference independent particle model.

the single-determinant representation cannot again describe correctly the true wavefunction as a single determinant.

To prove that determinants form an orthonormal basis for the multi-electronic system [102] we start with the expansion of a mono-electronic wavefunction in terms of an orthonormal and complete basis formed by spin-orbitals

$$\Psi(\boldsymbol{x}) = \sum_{k} C_k \psi_k(\boldsymbol{x}). \tag{1–22}$$

Repeating this procedure for all particles we see that, in general terms, an antisymmetric wavefunction is a linear combination of Hartree products like

$$\Psi(\underline{\boldsymbol{x}}) = \sum_{k_1,k_2,\ldots,k_N} C_{k_1,k_2,\ldots,k_N} \psi_{k_1}\psi_{k_2}\cdots\psi_{k_N}. \tag{1–23}$$

The coefficients of the expansion are antisymmetric in their indices, hence only a fraction of the coefficients are linearly independent: those that fulfill $k_1 < k_2 < \ldots < k_N$. A particular arrangement of this type is called an ordered configuration $K = (k_1, k_2, \ldots, k_N)$. Thus, the expansion in terms of the ordered configurations becomes an expansion in Slater determinants

$$\Psi(\underline{\boldsymbol{x}}) = \sum_{K} C_K \Psi_K, \qquad \Psi_K = |\psi_{k_1}\psi_{k_2}\cdots\psi_{k_N}\rangle. \tag{1–24}$$

Obviously, when there is only one configuration, this expansion is a determinantal function.

Deepening into the connection between its mathematical structure and its physical interpretation we can add more general conditions that the wavefunction has to satisfy when two electrons or an electron and a nucleus coalesce [87]:

$$\lim_{r_{ij}\to 0}\left(\frac{\partial\Psi}{\partial r_{ij}}\right)_{av.} = \frac{1}{2}\Psi(r_{ij}=0) \qquad r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \tag{1–25}$$

$$\lim_{r_{iA}\to 0}\left(\frac{\partial\Psi}{\partial r_{iA}}\right)_{av.} = -Z_A\Psi(r_{iA}=0) \qquad r_{iA} = |\mathbf{r}_i - \mathbf{R}_A|. \tag{1–26}$$

Based on the first cusp correlation condition the first order wavefunctions of helium singlet states in a perturbative treatment around the singularity have a dependence on the coordinate r_{12} that is

$$\Psi(\mathbf{r}_1,\mathbf{r}_2) = \Psi(r_{12}=0)\left\{1+\frac{1}{2}r_{12}\right\} + O(r_{12}^2). \tag{1–27}$$

Present methods based on expansions in a basis of determinants made of smooth orbitals have a very poor representation of the correlation cusp. Contrarily, trial functions with explicit dependence in the inverse distance of the electronic positions have a lightning rate of energy convergence as evidenced by Hylleraas in his landmark calculation of the ground state ionization potential of helium [83]. Sadly, this method is difficult to extend for larger systems.

1.3.1 The statistical picture: generalized wave functions

All the correlation is encoded in the exact wavefunction. Besides it is a complicated object that does not attend to particle indistinguishability without approximations. We can make $\|\Psi_{\text{exact}} - \Psi\|$ as small as we wish at the expense of a high computational cost but the orbital character is lost so it is not clear how correlation is affecting them. Instead, as is manifested by Hamiltonians, operators admit a one-, two-, three, ..., n-body expansion in Taylor series

$$O = O_0 + \sum_i O_i + \frac{1}{2!}\sum_{i\neq j} O_{ij} + \cdots . \tag{1–28}$$

Observations are this way separated by particle rank without violating the indistinguishability. Based on this finding if we encounter a state-operator correspondence we can separate the distribution of an electron (any) from the distribution of two electrons (any pair), and so on. There is, however, a natural (orthogonal) projector associated with Ψ that takes the role of a density operator. This establishes the nexus that allows us to set a hierarchy of electron correlations.

Density Matrices

Density matrices were firstly introduced to apply statistical concepts in quantum mechanics by John Von Neumann and Landau. When a physical system is at $T = 0K$ it may be understood as being isolated from the environment and its state may thus be represented by only one state vector (a pure state). At finite temperatures ($T \neq 0K$) thermal equilibrium will force it into a superposition of states (a mixed state) not defined by a single wavefunction pure state but by a weighted combination of them, or equivalently by a density matrix which describes the probability of finding the system into any of the available pure states. From this perspective density matrices generalize the concept of wavefunction. In our treatment of chemical systems we shall restrict ourselves to pure states. Notwithstanding, density matrices prove also to be useful under such circumstances.

The aforementioned density operator, Γ_N, acts on an arbitrary function Φ as a projector to the state Ψ

$$\Gamma_N|\Phi\rangle = \langle\Psi|\Phi\rangle|\Psi\rangle. \tag{1–29}$$

This Γ_N operator has many fundamental properties: idempotent $\Gamma^2\Psi = \Gamma\Psi$, self-adjoint $\Gamma = \Gamma^*$, positive semidefinite $\langle\Phi|\Gamma|\Phi\rangle \geqslant 0$, possessing one eigenvalue equal to one $\Gamma\Psi = \Psi$, and with unit trace $\mathrm{Tr}(\Gamma) = 1$.

Another important way to define an arbitrary linear operator, O, is by its kernel, K. The kernel K can be thought of as a generalization of the matrix

representation, M, of a linear map in finite dimensions, L, when the vectors are replaced by functions with values at infinitely many points:

$$(Lu)_i = \sum_j M_{ij} u_j \qquad \text{finite dim} \tag{1–30}$$

$$(O\Psi)(x) = \int K(x;x')\Psi(x')\,dx' \qquad \text{infinite dim.} \tag{1–31}$$

Abusing of notation, the density operator has a Green kernel $\Gamma(\underline{x};\underline{x}')$ as in equation (1–31) that keeps the equality (1–29)

$$\begin{aligned}(\Gamma_N\Phi)(\underline{x}) &= \int \Gamma(\underline{x};\underline{x}')\Phi(\underline{x}')\,d\underline{x}' = \int \left[\Psi(\underline{x})\Psi^*(\underline{x}')\right]\Phi(\underline{x}')\,d\underline{x}' \\ &= \left[\int \Psi^*(\underline{x}')\Phi(\underline{x}')\,d\underline{x}'\right]\Psi(\underline{x}).\end{aligned} \tag{1–32}$$

We used a contracted notation $\int d\underline{x}$ for $\sum_\sigma \int d\underline{r}$. Diagonal elements $\Gamma(\underline{x};\underline{x})$ are not defined by the previous equation. Otherwise the integration would also run over $\Psi(\underline{x})$. But we can define them directly as

$$\Gamma(\underline{x};\underline{x}) = \Psi(\underline{x})\Psi^*(\underline{x}) = |\Psi(\underline{x})|^2 \tag{1–33}$$

to have a formal definition of the trace

$$\mathrm{Tr}(\Gamma) = \int \Gamma(\underline{x};\underline{x})\,d\underline{x} = 1. \tag{1–34}$$

With this appreciation we are equipped to evaluate expectation values of multiplicative operators such as the potential from the density matrix

$$\langle V\rangle = \langle\Psi|V|\Psi\rangle = \int \Psi^*(\underline{x})V\Psi(\underline{x})\,d\underline{x} = \int V\Psi^*(\underline{x})\Psi(\underline{x})\,d\underline{x} = \int V\Gamma(\underline{x};\underline{x})\,d\underline{x} = \mathrm{Tr}(V\Gamma) \tag{1–35}$$

but not ready to compute — let's say — the kinetic energy, for the differential operator needs non diagonal elements to characterize the situation in other representations than the coordinate space. Just in case the operator is allowed only to act over unprimed coordinates at the same time that we prime the coordinates of $\Psi^*(\underline{x}')$ to avoid the action of the operator. Once the operator has done its job the primed coordinates are restored, $\underline{x}' = \underline{x}$, so

$$\mathrm{Tr}(T\Gamma) = \int T\Gamma(\underline{x};\underline{x})\,d\underline{x} = \int T\Psi(\underline{x})\Psi^*(\underline{x}')\,d\underline{x} = \int (T\Psi)(\underline{x})\Psi^*(\underline{x})\,d\underline{x} = \langle\Psi|T|\Psi\rangle = \langle T\rangle. \tag{1–36}$$

In this way, the functional that obtains the ground state energy from the wavefunction is mapped into a functional of the density matrix Γ:

$$\inf_\Gamma \left\{\,\mathrm{Tr}(H\Gamma)\right\} = \mathcal{E}(\Psi) = \inf_\Psi \langle\Psi|H|\Psi\rangle \tag{1–37}$$

Reduced Density Matrices (RDMs) and density functions

The complexity of the density matrix is on par with that of the wavefunction. For this reason Husimi proposed a reduction to k-particle internal distribution functions by integration of the density matrix:

$$\gamma_k(\mathbf{x}_{i\leqslant k};\mathbf{x}'_{i\leqslant k}) = \frac{N!}{(N-k)!}\int \Gamma(\mathbf{x}_{i\leqslant N};\mathbf{x}'_{i\leqslant k},\mathbf{x}_{k+1\leqslant i\leqslant N})\,d\mathbf{x}_{i>k}. \tag{1–38}$$

The reader should be aware that several normalization conventions exist [102, 126]. Indistinguishability of electrons imposes that any $N-k$ electrons could have been picked to be integrated out. All the electrons are equivalent so we can make an arbitrary choice of the variables to integrate and add a combinatoric prefactor that takes into account all possibilities. Reduced density matrices inherit from the wave function several fundamental properties, they are

$$\gamma_k(\mathbf{x}_{i\leqslant k};\mathbf{x}'_{i\leqslant k}) = \gamma^*_k(\mathbf{x}'_{i\leqslant k};\mathbf{x}_{i\leqslant k}) \qquad \text{hermitian,} \tag{1–39}$$

$$\gamma^*_k(\cdots,\mathbf{x}_i\cdots,\mathbf{x}_j,\cdots;\mathbf{x}'_{i\leqslant k}) = -\gamma^*_k(\cdots,\mathbf{x}_j\cdots,\mathbf{x}_i,\cdots;\mathbf{x}'_{i\leqslant k}) \qquad \text{antisymmetric,} \tag{1–40}$$

$$\gamma_k(\cdots,\mathbf{x}_i,\cdots,\mathbf{x}_i,\cdots) = 0 \qquad \text{showing exclusion,} \tag{1–41}$$

$$\mathrm{Tr}(\gamma_k) = \frac{N!}{(N-k)!} \qquad \text{finite trace,} \tag{1–42}$$

and positive semidefinite. Not only it is possible to obtain reduced density matrices from the full density matrix, there is indeed a recursive formula relating a p-density matrix with its lower reduced $(p-1)$-density matrices

$$\gamma_{p-1}(\mathbf{x}_{i\leqslant p-1};\mathbf{x}'_{i\leqslant p-1}) = \frac{p}{N-p+1}\int \gamma_p(\mathbf{x}_{i\leqslant p};\mathbf{x}'_{i\leqslant p-1},\mathbf{x}_p)\,d\mathbf{x}_p. \tag{1–43}$$

Each kernel has an associated operator that acts on functions of k variables by integration

$$(\gamma_k f_{1\ldots k})(\mathbf{x}_{i\leqslant k}) = \int \gamma_k(\mathbf{x}_{i\leqslant k};\mathbf{x}'_{i\leqslant k})\Psi(\mathbf{x}_{i\leqslant k})\,d\mathbf{x}_{i\leqslant k} \tag{1–44}$$

and returns a function of k variables. Since any wavefunction admits an expansion in products of N spin-orbitals, reduced density matrices should admit an expansion in k spin-orbitals

$$\gamma_k(\mathbf{x}_{i\leqslant k};\mathbf{x}'_{i\leqslant k}) = \sum_{\substack{u_1,\cdots,u_k\\ v_1,\cdots,v_k}} \gamma_k(u_1,\cdots,u_k|v_1,\cdots,v_k)\,\psi_{u_1}(\mathbf{x}_1)\cdots\psi_{u_k}(\mathbf{x}_k)\psi^*_{v_1}(\mathbf{x}'_1)\cdots\psi^*_{v_k}(\mathbf{x}'_k) \tag{1–45}$$

or in Slater determinants as in equation (1–24)

$$\gamma_k(\mathbf{x}_{i\leqslant k};\mathbf{x}'_{i\leqslant k}) = k! \sum_{\substack{u_1<\cdots<u_k \\ v_1<\cdots<v_k}} \gamma_k(u_1,\cdots,u_k|v_1,\cdots,v_k) \\ |\psi_{u_1}(\mathbf{x}_1)\cdots\psi_{u_k}(\mathbf{x}_k)\rangle\langle\psi^*_{v_1}(\mathbf{x}'_1)\cdots\psi^*_{v_k}(\mathbf{x}'_k)|. \tag{1–46}$$

Since most interesting operators in Quantum Mechanics are one- and two-electron operators, the most important DMs are the first and second rank ones. Diagonalization of γ_1 yields an eigenbasis of one particle natural spin-orbitals [102] with eigenvalues that denote their natural occupation. The occupations lie in the range $[0,1]$ and are normalized to $\mathrm{Tr}\,\gamma_1 = N = \sum_i n_i$. The second order density matrix has an eigenbasis of geminals with occupations that comply with equivalent conditions.

Reduced density matrices are not only practical entities, they contain all the relevant information about the physical system, for example

$$\rho(\mathbf{x}) = \gamma_1(\mathbf{x},\mathbf{x}) := \sum_k^N \int |\Psi(\underline{\mathbf{x}})|^2 \, d\mathbf{x}_{i\neq k} \tag{1–47}$$

$$\rho_2(\mathbf{x}_1,\mathbf{x}_2) = \gamma_2(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1,\mathbf{x}_2) \tag{1–48}$$

are the spin (one-electron) density and the spin (electron) pair density.

In summary we have established a map from the wavefunction to the density matrix, then we have narrowed down the number of particles described, to finally represent density matrices in position space through their diagonals as probability density functions:

$$\Psi \mapsto \Gamma \mapsto \cdots \gamma_{1\leqslant p\leqslant N} \cdots \mapsto \rho_p. \tag{1–49}$$

Only remains to us the explicit evaluation of density matrices. We will look first at an easy example to capture their structure.

Example 1.1 (The case of a Slater determinant)**.** *A system with $N = 2$ electrons in a configuration $\{\psi_1^1, \psi_2^1\}$ has a wavefunction*

$$\Psi = \frac{1}{\sqrt{2!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) \end{vmatrix}. \tag{1–50}$$

Its first order density matrix is

$$\gamma_1(\mathbf{x}_1;\mathbf{x}'_1) = \int \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) \end{vmatrix} \cdot \begin{vmatrix} \psi_1^*(\mathbf{x}'_1) & \psi_2^*(\mathbf{x}'_1) \\ \psi_1^*(\mathbf{x}_2) & \psi_2^*(\mathbf{x}_2) \end{vmatrix} d\mathbf{x}_2 \\ = \left[\psi_1(\mathbf{x}_1)\psi_1^*(\mathbf{x}'_1) + \psi_2(\mathbf{x}_1)\psi_2^*(\mathbf{x}'_1)\right] = \sum_i^N \psi_i(\mathbf{x}_1)\psi_i^*(\mathbf{x}'_1) \tag{1–51}$$

In return $\gamma_1(\mathbf{x}_1;\mathbf{x}'_1) \equiv \gamma_1(\mathbf{x}_2;\mathbf{x}'_2)$. The last equality resembles the closure relation $\sum_i^\infty \psi_i(\mathbf{x}_1)\psi_i^*(\mathbf{x}'_1) = \delta(\mathbf{x}_1 - \mathbf{x}'_1)$ that stems from a complete basis for the mono electronic space. Any one electron function is $f_1(\mathbf{x}) = \sum_i^N c_i\psi_i(\mathbf{x})$. Spinorbitals constitute an eigenbasis for the one-particle density operator with eigenvalues equal to unity

$n_1 = n_2 = 1$ *for the orbitals* ψ_1 *and* ψ_2*. This is evident from the beginning since we stated that the unique configuration is* $\{\psi_1^1, \psi_2^1\}$*. That is not necessarily true for correlated wavefunctions. These orbitals that diagonalize the first order density matrix are called natural spin orbitals and their eigenvalues natural occupations. The second order density matrix, or full density matrix (*$k = N$*), needs no integration of* $\Psi\Psi^*$

$$\begin{aligned}\gamma_2(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1',\mathbf{x}_2') &= \begin{vmatrix}\psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1)\\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2)\end{vmatrix}\cdot\begin{vmatrix}\psi_1^*(\mathbf{x}_1') & \psi_2^*(\mathbf{x}_1')\\ \psi_1^*(\mathbf{x}_2') & \psi_2^*(\mathbf{x}_2')\end{vmatrix}\\ &= \psi_1(\mathbf{x}_1)\psi_1^*(\mathbf{x}_1')\psi_2(\mathbf{x}_2)\psi_2^*(\mathbf{x}_2') - \psi_1(\mathbf{x}_1)\psi_1^*(\mathbf{x}_2')\psi_2(\mathbf{x}_2)\psi_2^*(\mathbf{x}_1')\\ &- \psi_1(\mathbf{x}_2)\psi_1^*(\mathbf{x}_1')\psi_2(\mathbf{x}_1)\psi_2^*(\mathbf{x}_2') + \psi_1(\mathbf{x}_2)\psi_1^*(\mathbf{x}_2')\psi_2(\mathbf{x}_1)\psi_2^*(\mathbf{x}_1')\\ &= \sum_{k\neq l}^{N}\psi_k(\mathbf{x}_1)\psi_k^*(\mathbf{x}_1')\psi_l(\mathbf{x}_2)\psi_l^*(\mathbf{x}_2') - \psi_k(\mathbf{x}_1)\psi_k^*(\mathbf{x}_2')\psi_l(\mathbf{x}_2)\psi_l^*(\mathbf{x}_1')\end{aligned}$$

Identifying the first and fourth terms as part of $\gamma_1(\mathbf{x}_1;\mathbf{x}_1')\gamma_1(\mathbf{x}_2;\mathbf{x}_2')$ *from (1–51) and subtracting the missing terms*

$$\begin{aligned}\gamma_2(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1',\mathbf{x}_2') &= \gamma_1(\mathbf{x}_1;\mathbf{x}_1')\gamma_1(\mathbf{x}_2;\mathbf{x}_2') - \sum_{k,l}\psi_k(\mathbf{x}_1)\psi_k^*(\mathbf{x}_2')\psi_l(\mathbf{x}_2)\psi_l^*(\mathbf{x}_1')\\ &= \gamma_1(\mathbf{x}_1;\mathbf{x}_1')\gamma_1(\mathbf{x}_2;\mathbf{x}_2') - \gamma_1(\mathbf{x}_1;\mathbf{x}_2')\gamma_1(\mathbf{x}_2;\mathbf{x}_1')\\ &= \begin{vmatrix}\gamma_1(\mathbf{x}_1;\mathbf{x}_1') & \gamma_1(\mathbf{x}_1;\mathbf{x}_2')\\ \gamma_1(\mathbf{x}_2;\mathbf{x}_1') & \gamma_1(\mathbf{x}_2;\mathbf{x}_2')\end{vmatrix}\end{aligned} \tag{1–52}$$

Surprisingly, in single determinant cases, high order density matrices can be recovered from a lower rank one by using the above expansion, which is called the Fock-Dirac expansion. Using induction, higher order matrices become

$$\gamma_k(\mathbf{x}_{i\leqslant k};\mathbf{x}'_{i\leqslant k}) = k!\sum_{(i_1,\ldots,i_k)\subseteq[1,N]}|\psi_{i_1}(\mathbf{x}_1)\cdots\psi_{i_k}(\mathbf{x}_k)\rangle\langle\psi_{i_1}(\mathbf{x}_1')\cdots\psi_{i_k}(\mathbf{x}_k')|. \tag{1–53}$$

The diagonal part of the 1-RDM is the density function

$$\rho(\mathbf{x}_1) = \sum_{i}^{N}\psi_i(\mathbf{x}_1)\psi_i^*(\mathbf{x}_1) = \sum_{i}^{N}|\psi_i(\mathbf{x}_1)|^2 \tag{1–54}$$

and the diagonal of 2-RDM is the pair density function

$$\begin{aligned}\rho_2(\mathbf{x}_1,\mathbf{x}_2) &= \rho(\mathbf{x}_1)\rho(\mathbf{x}_2) - |\gamma_1(\mathbf{x}_1;\mathbf{x}_2)|^2\\ &= \sum_{k\neq l}^{N}|\psi_k(\mathbf{x}_1)|^2|\psi_l(\mathbf{x}_2)|^2 - \psi_k(\mathbf{x}_1)\psi_k^*(\mathbf{x}_2)\psi_l(\mathbf{x}_2)\psi_l^*(\mathbf{x}_1)\end{aligned} \tag{1–55}$$

follows from the hermitian property $\gamma_1(\mathbf{x}_1;\mathbf{x}_2) = \gamma_1^*(\mathbf{x}_2;\mathbf{x}_1)$*. Both equalities come from different paths but they are equivalent. To see the link notice that in the last sum we can eliminate the restriction of self pairing* $k \neq l$ *because the first and second term cancel* $\sum_k |\psi_k(\mathbf{x}_1)|^2|\psi_k(\mathbf{x}_2)|^2 - |\psi_k(\mathbf{x}_1)|^2|\psi_k(\mathbf{x}_2)|^2$*. The first term converts into a*

product of independent probability densities with contribution $\mathrm{Tr}(\rho(\mathbf{x}_1)\rho(\mathbf{x}_2)) = N^2$, *instead of* $N(N-1)$ *as before, that contains part of the direct product and part of the exchange contribution. The second an "exchange" term with the rest of contributions of the direct and exchange products that has a trace of* $\mathrm{Tr}(-|\gamma_1(\mathbf{x}_1;\mathbf{x}_2)|^2) = -N$, *not* 0 *as before. The negative correlation* $-|\gamma_1(\mathbf{x}_1;\mathbf{x}_2)|^2$ *or probability reduction of the joint probability* $\rho_2(\mathbf{x}_1,\mathbf{x}_2)$ *with respect to the product, that is named Fermi correlation, is more due to avoidance of self pairing (Pauli correlation) than to antisymmetry. Already in this example density matrices and density functions permit us appreciate the subtle origin for the departure from the uncorrelated motion. If* ψ_1 *and* ψ_2 *share the same spatial orbitals (closed shell), (i) the density is* $\rho(\mathbf{x}) = 2\sum_i^{N/2}|\phi_i(\mathbf{r})|^2$ *and (ii) in this special case of two electrons alone, the second term of* (1–55) *cancels resulting in an almost uncorrelated motion. Density matrices of determinantal functions have only diagonal elements populated. As we increase correlation the non-diagonal elements become populated.*

To make the partition of observables by particle rank more precise, let us consider the Taylor expansion of the density operator, symmetric in all indices, as a sum of projections over single electron functions f_i, geminals (pair products of spin-orbitals) $f_{ij} = \psi_i\psi_j, \ldots$

$$\begin{aligned}\Gamma_N &= \Gamma_0 + \sum_i^N |f_i\rangle\langle f_i| + \frac{1}{2!}\sum_{i\neq j}^N |f_{ij}\rangle\langle f_{ij}| + \cdots .\\ &= \Gamma_0 + \sum_i^N \gamma_1(i) + \frac{1}{2!}\sum_{i\neq j}^N \gamma_2(ij) + \cdots .\\ &= \Gamma_0 + N\gamma_1 + \binom{N}{2}\gamma_2 + \cdots . \end{aligned} \tag{1–56}$$

Using reduced density matrices the expectation of the two particle projector is

$$\begin{aligned}\binom{N}{2}\langle\Psi|\gamma_2|\Psi\rangle &= \frac{1}{2!}\int \Psi^*\gamma_2(\mathbf{x}_{i<2};\mathbf{x}'_{i<2})\Psi\,d\underline{\mathbf{x}}\\ &= \frac{1}{2}\int f^*_{12}(\mathbf{x}_{i<2})\gamma_2(\mathbf{x}_{i<2};\mathbf{x}'_{i<2})f_{12}(\mathbf{x}'_{i<2})\,d\mathbf{x}_1\,d\mathbf{x}_2\\ &= \frac{1}{2}f^*_{12}(\mathbf{x}_{i<2})f_{12}(\mathbf{x}_{i<2}) = \frac{1}{2}|f_{12}(\mathbf{x}_{i<2})|^2 \end{aligned} \tag{1–57}$$

a two-electron distribution function. We have operators acting on a space that is a direct sum of single particle tensor powers spaces with implicit antisymmetrization.

$$|\Psi\rangle = |f_1\rangle|f_{12}\rangle|f_{123}\rangle\cdots . \tag{1–58}$$

Any physical quantity involving k-particle processes can be stated in terms of the corresponding k-density matrix. For example, the energy of a general Hamiltonian involving one-particle operators O_1 and two-particle operators O_2 is

$$\mathcal{E}(\Gamma) = \mathrm{Tr}\,H\Gamma = \mathrm{Tr}\,O_1\gamma_1 + \frac{1}{2}\mathrm{Tr}\,O_2\gamma_2. \tag{1–59}$$

This leads us to think that we only need the second order reduced density matrix to compute the ground state energy instead of all N-particle distributions. The unique needed element would be γ_2, from γ_2 the one particle density matrix is straightforward $\gamma_1 = \frac{2}{N-1} \mathrm{Tr}^1 \gamma_2$. In this sense the wavefunction or the density matrix contain too much information, they are over–complete.

An appealing feature of the 2RDM is that all interactions can be correctly represented so the energy functional is completely known. It might seem that nothing has been lost in the transition from the N particle wavefunction to the 2RDM, though we have missed the Hilbert space boundary conditions or more commonly called N-representability [36] conditions. A variational computation of the wavefunction is carried out minimizing the energy with respect to variations of the trial function so they span the whole space of functions. In any case those variations are never arbitrary, square integrability and regularity conditions must be satisfied, that way the wavefunction has the properties of a distribution of electrons. The 2-RDM needs additional constraints, apart from being integrated from the density matrix to compute the energy of the ground state: and admissible conditions. For a long time, many necessary but not sufficient or impractical sufficient conditions were known. It has been only quite recently that full N-representability conditions have been identified, see [124, 125] for an account of the latest developments, and partially implemented in polynomial time [123] for realistic systems.

Cumulants and correlation holes

Let us realize that the density functions are the moments of the electronic distribution. If a generating function for the moments ρ_r exists $M(\zeta) = \langle e^{\zeta x} \rangle$ there is an alternative generating function for the density cumulants $\ln M(\zeta)$ that when expanded as a Maclaurin series gives the cumulants κ_r as r times the derivative of its generating function at $\zeta = 0$

$$\kappa_r = \frac{d^r \ln M(\zeta)}{d\zeta^r}\Big|_{\zeta=0} = \sum_{i=1}^{r} (-1)^{i-1} (i-1)! B_{r,i}(\rho_1, \cdots, \rho_{r-i+1}) \qquad (1\text{–}60)$$

With the aid of Bell polynomials $B_r(\rho_1,\dots,\rho_r)=\sum_{i=1}^{r} B_{r,i}(\rho_1,\rho_2,\dots,\rho_{r-i+1})$

$$B_r(\rho_1,\dots,\rho_r)=\begin{vmatrix} \rho_1 & \binom{r-1}{1}\rho_2 & \binom{r-1}{2}\rho_3 & \binom{r-1}{3}\rho_4 & \binom{r-1}{4}\rho_5 & \cdots & \cdots & \rho_r \\ -1 & \rho_1 & \binom{r-2}{1}\rho_2 & \binom{r-2}{2}\rho_3 & \binom{r-2}{3}\rho_4 & \cdots & \cdots & \rho_{r-1} \\ 0 & -1 & \rho_1 & \binom{r-3}{1}\rho_2 & \binom{r-3}{2}\rho_3 & \cdots & \cdots & \rho_{r-2} \\ 0 & 0 & -1 & \rho_1 & \binom{r-4}{1}\rho_2 & \cdots & \cdots & \rho_{r-3} \\ 0 & 0 & 0 & -1 & \rho_1 & \cdots & \cdots & \rho_{r-4} \\ 0 & 0 & 0 & 0 & -1 & \cdots & \cdots & \rho_{r-5} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & -1 & \rho_1 \end{vmatrix} \tag{1–61}$$

we obtain the usual cumulant densities

$$\kappa_1=\rho_1 \tag{1–62}$$

$$\kappa_2=\rho_2-{\rho_1}^2 \tag{1–63}$$

$$\kappa_3=\rho_3-3\rho_2\rho_1+2{\rho_1}^3 \tag{1–64}$$

$$\kappa_4=\rho_4-4\rho_3\rho_1-3{\rho_2}^2+12\rho_2{\rho_1}^2-6{\rho_1}^4. \tag{1–65}$$

The reader might question what the purpose for introducing this new formalism is. One of the reasons is that k density functions quantify the probability of k electrons hitting each other simultaneously as well as $k-1, k-2, \cdots, 2$ collisions, nested like a Matryoshka. They include all possible partitions of k correlation. Cumulants, on the opposite hand, only have genuine k correlations. The first order cumulant is the difference between an electron distribution and one electron correlation (not existing), the second order cumulant is the pair electron density function minus the distribution of two independent electrons, the third one is the triad distribution minus non equivalent pair correlations plus non correlated motion. Cumulant densities implement a true correlation ladder increasing correlation progressively. They are Hermitian and antisymmetric as the density functions and have the remarkable property of being additive for independent variables, the first step to be extensive. The trace of the cumulants is proportional to the size of the system $\mathcal{O}(N)$:

$$\mathrm{Tr}(\kappa_1)=N \qquad \mathrm{Tr}(\kappa_2)=-N \tag{1–66}$$

$$\mathrm{Tr}(\kappa_3)=2N \qquad \mathrm{Tr}(\kappa_4)=-6N \tag{1–67}$$

For future purposes a convenient change of definition is due

1. Divide all terms of κ_r by the coefficient that is in front of ρ_1^r.
2. Multiply each term $\rho_{s_1}^{p_1}\rho_{s_2}^{p_2}\cdots\rho_{s_k}^{p_k}$ by $\frac{\prod_{i=1}^{k} p_i!(s_i!)^{p_i}}{l!}$.

$\rho(\mathbf{x}) = \rho_1(\mathbf{x})$ introduces a new set of cumulants

$$\lambda_1(\mathbf{x}) = \gamma_1(\mathbf{x};\mathbf{x}) = \rho(\mathbf{x}) \tag{1–68}$$

$$\lambda_2(\mathbf{x}_1,\mathbf{x}_2) = \rho(\mathbf{x}_1)\rho(\mathbf{x}_2) - \rho_2(\mathbf{x}_1,\mathbf{x}_2) \tag{1–69}$$

$$\begin{aligned}\lambda_3(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) = {}& \rho(\mathbf{x}_1)\rho(\mathbf{x}_2)\rho(\mathbf{x}_3) \\ & - \frac{1}{2}\left[\rho(\mathbf{x}_1)\rho_2(\mathbf{x}_2,\mathbf{x}_3) + \rho(\mathbf{x}_2)\rho_2(\mathbf{x}_1,\mathbf{x}_3) + \rho(\mathbf{x}_3)\rho_2(\mathbf{x}_1,\mathbf{x}_2)\right] \\ & + \frac{1}{2}\rho_3(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)\end{aligned} \tag{1–70}$$

with two very important properties, namely recursivity

$$\lambda_{k-1}(\mathbf{x}_{i\leqslant k-1}) = \int \lambda_k(\mathbf{x}_{i\leqslant k})\,d\mathbf{x}_k \tag{1–71}$$

and extensivity

$$\mathrm{Tr}(\lambda_k) = \int \lambda_k(\mathbf{x}_{i\leqslant k})\,d\mathbf{x}_{i\leqslant k} = N. \tag{1–72}$$

The second cumulant is also referred as the exchange-correlation pair density alluding to the Fermi (exchange) and Coulomb correlation of pairs of electrons that it integrates and represents the part that cannot be recovered from the density. In DFT (section 1.7) it is a most valuable quantity. Nonetheless, it is also appropriate to pay careful attention to its statistical meaning because we could encounter surprises. A system with two independent electrons and normalized densities has a normalized joint probability distribution

$$\rho_2(\mathbf{x}_1,\mathbf{x}_2) = \frac{n-1}{n}\rho(\mathbf{x}_1)\rho(\mathbf{x}_2). \tag{1–73}$$

Even for a non correlated system the pair density is not just the product of the individual densities by a factor $(-1/n)\rho(\mathbf{x}_1)\rho(\mathbf{x}_2)$. It results practical to focus on the distribution of one electron to circumvent the normalization. It would be a conditional probability to encounter an electron at $\mathbf{x}_2$ given the event that another electron was found at $\mathbf{x}_1$

$$\mathcal{P}(\mathbf{x}_2|\mathbf{x}_1) = \frac{\rho_2(\mathbf{x}_1,\mathbf{x}_2)}{\rho(\mathbf{x}_1)}, \tag{1–74}$$

that quantifies the exchange-correlation charge density h_{xc} by comparison with the independent density

$$\mathcal{P}(\mathbf{x}_2|\mathbf{x}_1) = \rho(\mathbf{x}_2) + h_{xc}(\mathbf{x}_1,\mathbf{x}_2). \tag{1–75}$$

The probability to encounter an electron in the whole space after fixing the position of one of them at $\mathbf{x}_1$ is

$$\int \mathcal{P}(\mathbf{x}_2|\mathbf{x}_1)\,d\mathbf{x}_2 = N - 1. \tag{1–76}$$

$h_{xc}(\mathbf{x}_1,\mathbf{x}_2)$ integrates to minus one over the space of electron 2, hence the name of exchange-correlation hole. The effect of exchange alone can be isolated taking a look at single determinant wavefunctions, the *Fermi hole*

$$h_x(\mathbf{x}_1,\mathbf{x}_2) = -\frac{|\gamma_1(\mathbf{x}_1;\mathbf{x}_2)|^2}{\rho(\mathbf{x}_1)} = -\sum_{k,l}^{N} \frac{\psi_k(\mathbf{x}_1)\psi_k^*(\mathbf{x}_2)\psi_l(\mathbf{x}_2)\psi_l^*(\mathbf{x}_1)}{|\psi_k(\mathbf{x}_1)|^2} \tag{1–77}$$

also integrates to minus one. Suppose that the fixed electron at $\mathbf{r}_1$ is well localized in an orbital $\psi_k(\mathbf{x}_1)$, the excluded density is the occupation of the same orbital at $\mathbf{x}_2$

$$h_x(\mathbf{x}_1,\mathbf{x}_2) = -|\psi_k(\mathbf{x}_1)|^2. \tag{1–78}$$

From this expression we can appreciate once again that the source of the exchange-only hole is not always Fermi statistics, it can be the avoidance of self pairing in a shell structure. In any case the center of the Fermi hole is

$$\lim_{\mathbf{r}_2\to\mathbf{r}_1} h_x(\mathbf{x}_1,\mathbf{x}_2) = -\sum_{l}^{N} |\psi_l(\mathbf{x}_2)|^2 = -\rho(\mathbf{x}_2). \tag{1–79}$$

The missing Coulomb correlation is expected to contribute nothing to the trace. On average it evicts and concentrates charge with no cumulative effect. Since the origin is the inter-electronic Hamiltonian operator g_{ij} and no restriction on the wavefunction the *Coulomb hole* is expected to take place only in the vicinity of the other electrons with a counter effect, a *Coulomb heap*, in the rest of the space.

We have seen that the second order density function is recoverable at the single determinantal level from the first order density matrix and we advance that any density function of order $2 \leqslant k \leqslant N$

$$\rho_k(\mathbf{x}_{i\leqslant k}) = \sum_{\tau} (-1)^p \gamma_1(\mathbf{x}_1;\mathbf{x}_{\tau_1}) \cdots \gamma_1(\mathbf{x}_N;\mathbf{x}_{\tau_N}) \tag{1–80}$$

is a sum of permutations τ of the diagonal first order density matrices with sign p each. For example,

$$\begin{aligned}
\rho_3(\mathbf{x}_{i\leqslant 3}) &= \gamma_1(\mathbf{x}_1;\mathbf{x}_1)\gamma_1(\mathbf{x}_2;\mathbf{x}_2)\gamma_1(\mathbf{x}_3;\mathbf{x}_3) + \gamma_1(\mathbf{x}_1;\mathbf{x}_1)\gamma_1(\mathbf{x}_2;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_2) \\
&+ \gamma_1(\mathbf{x}_2;\mathbf{x}_2)\gamma_1(\mathbf{x}_1;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_1) + \gamma_1(\mathbf{x}_3;\mathbf{x}_3)\gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1) \\
&+ \gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_1) + \gamma_1(\mathbf{x}_1;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1) \\
&= \rho(\mathbf{x}_1)\rho(\mathbf{x}_2)\rho(\mathbf{x}_3) + \rho(\mathbf{x}_1)\gamma_1(\mathbf{x}_2;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_2) \\
&+ \rho(\mathbf{x}_2)\gamma_1(\mathbf{x}_1;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_1) + \rho(\mathbf{x}_3)\gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1) \\
&+ \gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_1) + \gamma_1(\mathbf{x}_1;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1)
\end{aligned} \tag{1–81}$$

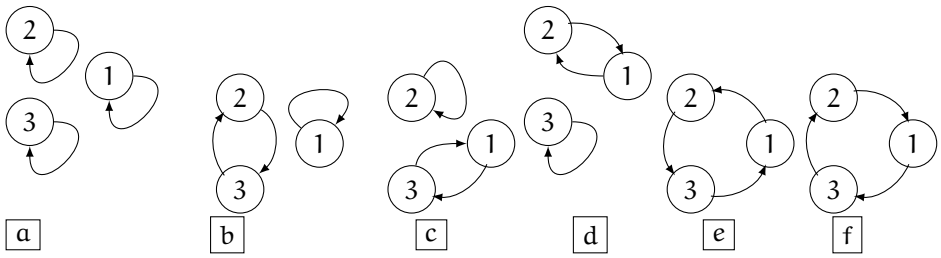


Figure 1: Diagrammatic partitions of three body correlations. (a) Disconnected one body correlations $\gamma_1(\mathbf{x}_1;\mathbf{x}_1)\gamma_1(\mathbf{x}_2;\mathbf{x}_2)\gamma_1(\mathbf{x}_3;\mathbf{x}_3)$ (b),(c),(d) Pair correlations $\gamma_1(\mathbf{x}_1;\mathbf{x}_1)\gamma_1(\mathbf{x}_2;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_2)$, $\gamma_1(\mathbf{x}_2;\mathbf{x}_2)\gamma_1(\mathbf{x}_1;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_1)$, $\gamma_1(\mathbf{x}_3;\mathbf{x}_3)\gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1)$ (e) Clockwise connected three body correlation $\gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_1)$ (f) Counter-clockwise connected three body correlation $\gamma_1(\mathbf{x}_1;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1)$

The first term is independent motion, the next three are pair correlations and the last two are identical three body correlations (figure 1). Consequently, the cumulant of third order is the average of the last two terms

$$\lambda_3 = \frac{1}{2}\left[\gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_1) + \gamma_1(\mathbf{x}_1;\mathbf{x}_3)\gamma_1(\mathbf{x}_3;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1)\right] \tag{1–82}$$

1.4 HARTREE–FOCK THEORY

We have already discussed about the algebraic structure of the wavefunction and the evaluation of the expected values at length but we have not debated how to obtain the optimal wavefunction. We know the exact expression for the Hamiltonian from (1–15) and the conditions to consider a many electron wavefunction admissible from the previous section but the orbitals have not been determined yet. The Hartree-Fock solution is that one that restricts variationally the multi–electronic wavefunction to a single Slater determinant (1–21). We are faced with a minimization problem where the Hartree-Fock energy

$$E^{HF} = \inf_{\Psi\in\mathcal{S}} \mathcal{E}^{HF}[\Psi] = \inf_{\Psi\in\mathcal{S}} \langle\Psi|H|\Psi\rangle \geqslant E_{gs}. \tag{1–83}$$

is always an upper bound for the ground state energy.

The energy functional $\mathcal{E}[\Psi]$ adopts a very compact expression on single determinant wavefunctions [37, 175]:

$$\begin{aligned}\mathcal{E}^{HF}[\Psi] &= \sum_i^N \langle\psi_i|h_i|\psi_i\rangle + \frac{1}{2}\sum_i^N\sum_{j\neq i}^N \left\{\langle\psi_i\psi_j|g_{ij}|\psi_i\psi_j\rangle - \langle\psi_i\psi_j|g_{ij}|\psi_j\psi_i\rangle\right\} \\ &= \sum_i^N \langle\psi_i|h_i|\psi_i\rangle - \frac{1}{2}\sum_i^N\sum_{j\neq i}^N \left\{J_{ij} - K_{ij}\right\}.\end{aligned} \tag{1–84}$$

Single particle core Hamiltonian operators h_i are restricted by orthonormality of the spin-orbitals to operate on their respective single-particle spaces. Otherwise, the bi-electronic operators are affected by the antisymmetry of the wavefunction under exchange of pairs. Two terms arise from their action, [4] one is the Coulomb integral J_{ij}, analogous to the classical Coulomb energy and a purely quantum exchange integral K_{ij}.

The entry point to model fermionic systems is to reckon the interacting system as a mean field where each orbital feels the average potential of the other electrons. That effect would be reached attaching an effective charge to the bi-electronic operators to form a new couple of single particle effective operators that replace the role of the Coulomb operator

$$J_j\psi_i = \langle\psi_j|g_{ij}|\psi_j\rangle\psi_i \tag{1–85}$$

and the exchange operator

$$K_j\psi_i = \langle\psi_j|g_{ij}|\psi_i\rangle\psi_j. \tag{1–86}$$

The Coulomb operator simulates the effect of the Coulomb potential on an electron at x_1 in the presence of another electron with density $|\psi_j(x_2)|^2$. The exchange operator interchanges the orbital in its kernel with its spinorbital argument. The Coulomb operator is local for it needs only the value of the orbital in the point where it is evaluated while the exchange operator is not local, needing the knowledge of the spinorbital it acts upon in the entire space. The effective Fock operator F is equal for all spinorbitals

$$\mathcal{E}^{\mathrm{HF}}[\Psi] = \sum_i \langle\psi_i|h_i + \sum_j (J_j - K_j)|\psi_i\rangle = \sum_i \langle\psi_i|F|\psi_i\rangle. \tag{1–87}$$

We do not restrict the sum over the index j for the same reason discussed in example 1.1.

Invoking the variational principle the best orbitals are those that minimize the energy functional, thus an arbitrary infinitesimal variation of the orbitals ($\psi_i \to \psi_i + \delta\psi_i$) should leave the energy unchanged. In addition, the energy has to be minimized under othonormalization constraints, that is to minimize

$$\mathcal{L}^{\mathrm{HF}}[\Psi] = \mathcal{E}^{\mathrm{HF}}[\Psi] - \sum_{ij} \lambda_{ij}\left(\langle\psi_i|\psi_j\rangle - \delta_{ij}\right). \tag{1–88}$$

4 Two different notations are present in the literature:

- physicists' notation:

$$\langle ij|kl\rangle = \langle\psi_i\psi_j|\psi_k\psi_l\rangle = \int \psi_i^*(x_1)\psi_j^*(x_2)\psi_k(x_1)\psi_l(x_2)\,dx_1\,dx_2$$

- chemists' notation:

$$[ij|kl] = [\psi_i\psi_j|\psi_k\psi_l] = \int \psi_i^*(x_1)\psi_j(x_1)\psi_k^*(x_2)\psi_l(x_2)\,dx_1\,dx_2$$

At a stationary point of the energy

$$\delta\mathcal{L}^{\mathrm{HF}}[\Psi] = \sum_i \left[\langle\psi_i|F|\delta\psi_i\rangle + \langle\psi_i|F|\delta\psi_i\rangle^* - \sum_j \left(\lambda_{ij}\langle\psi_i|\delta\psi_j\rangle + \lambda_{ij}\langle\psi_i|\delta\psi_j\rangle^* \right) \right]$$
$$= \sum_i \left[\langle\psi_i|F - \sum_j \lambda_{ij}|\delta\psi_i\rangle + \langle\psi_i|F - \sum_j \lambda_{ij}|\delta\psi_i\rangle^* \right] = 0. \tag{1–89}$$

Both the complex and the real solutions are equal and minimize the energy so we focus in the real part. Apart from the trivial solution to this system of N coupled differential equations where the spin-orbitals vanish (it has no physical meaning and does not satisfy normalization $\Psi = 0$) the inner part within the integration has to vanish. The wavefunction, a single Slater determinant, is invariant under unitary transformations U (e.g. permutation of columns), so the Fock operator and the Lagrange matrix are hermitian and the energy will remain invariant if we apply such transformation $E(\Psi U) = E(\Psi)$. The canonical equations that the best molecular orbitals must satisfy can be succinctly written in matrix notation

$$F(\psi)\psi_i = \varepsilon_i\psi_i \tag{1–90}$$

as a non-linear eigenvalue problem. The parenthesis is included to indicate that a knowledge of the Fock operator requires the knowledge of the spinorbitals, thus the solution is found iteratively in what is known as Self Consistent Field (SCF) procedure. The diagonalized matrix of Lagrange multipliers ε represents the one particle or spinorbital energies. To a good approximation orbital energies can be used estimate the ionization potential. While the sum over the occupied state energies is the energy of the system the sum over all minus the highest occupied orbital energy (HOMO) is roughly the energy of a $N-1$ system [91]. Increasing ionization levels are worse represented because the dismissed electronic cloud relaxation is more important.

Although our discussion has been centered around ground states, by the Hylleraas–Undheim–MacDonald theorem [84, 103], all excited states are readily available by simply forcing the functions of the variational space to be orthogonal to the low lying states.

Remember that we still don't know how those molecular orbitals are, we will only mention by now that the Hartree-Fock limit is achieved when the spinorbitals form a complete basis for the mono-electronic functional space.

A formulation of the HF equations in terms of reduced density matrices is

$$\mathcal{E}^{\mathrm{HF}}(\Gamma) = \mathrm{Tr}\, H\Gamma = \mathrm{Tr}\, h\gamma_1 + \frac{1}{2}\mathrm{Tr}\, g\rho_2$$
$$= \frac{1}{2}\int |\nabla\gamma_1(\mathbf{x};\mathbf{x}')|^2\, d\mathbf{x}\, d\mathbf{x}' - \sum_A \int \frac{\rho(\mathbf{x})Z_A}{|\mathbf{r}-\mathbf{R}_A|}\, d\mathbf{x} + \frac{1}{2}\int \frac{\rho(\mathbf{x})\rho(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{x}\, d\mathbf{x}'$$
$$- \frac{1}{2}\int \frac{|\gamma_1(\mathbf{x},\mathbf{x}')|^2}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{x}\, d\mathbf{x}'. \tag{1–91}$$

Only an off-diagonal operator (like the Laplacian) can feel the spin and symmetry properties of a wave function. The kinetic energy functional depends on the first-order density matrix, γ_1, the reason why the kinetic energy cannot be expressed as a function of the density alone is because it is not the sum of independent particle kinetic energies, as we stated below the electrons are correlated by the Pauli principle. The external potential (multiplicative) depends on the density, $\rho(\mathbf{x})$, and the electron interaction on the pair density, $\rho_2(\mathbf{x}, \mathbf{x}')$. The multiplicative nature of the Coulombic potential allows to rearrange the terms of the integrand making the dependence on the primed variables vanish. The kinetic energy is the only term that requires information not contained in the density or pair density. This observation becomes very relevant when discussing approximations for density functionals in the context of DFT.

There are three subordinate implementations of HF:

RESTRICTED HARTREE–FOCK (RHF) The wavefunction is restricted to have each spatial orbital doubly occupied (closed shell). It is valid only for systems with an even number of electrons. The total spin state is a singlet. Dissociation energies are wrong, both covalent and ionic structures have the same weight. Example of H_2 failure.

$$F = h_i + \sum_j^{N/2} 2J_j - K_j \tag{1–92}$$

UNRESTRICTED HARTREE–FOCK (UHF) The description of dissociation energies is better with the disadvantage of not being an eigenfunction of S^2. The Hamiltonian has no spin dependence, thus it should commute with spin operators $[H, S_z] = 0$, $[H, S^2] = 0$ and share the same eigenstate basis. Radical species.

RESTRICTED OPEN HARTREE–FOCK (ROHF) Alleviates the symmetry breaking problem adding those linear combinations of the uncoupled spinorbitals that result from permuting only the spin part. It is a multi determinant wavefunction.

Despite being a naive model, RHF delivers a very good approximation to the total energy of a system near to equilibrium. The energy that has not been recovered is the price that we pay for simplifying an interacting system by a mean field. The RHF solution thereupon will become a basis for later methods that aim to improve the description of correlations and a reference for Coulomb correlation measures. For closed-shell states

$$E_{corr} = E_{gs} - E^{HF} \tag{1–93}$$

is widely used although it lacks the insight that density matrices provide. Correlation energies have very tiny magnitudes compared to the total energy — $\approx 1\%$ — but are essential to obtain molecular properties accurate enough to guide future experiments. However, the relevant quantities are the energetic

differences among chemical compounds rather than the absolute energy of a system. It has been found that reaction energies predicted with HF are in qualitative agreement with experiments. Errors committed for one system are replicated for other similar system. This reflects an attracting feature of theoretical models: errors are almost systematic.

1.5 DISCRETIZATION OF THE HILBERT SPACE WITH ONE-ELECTRON BASIS SETS

Up to this time, we have said nothing about the shape of the orbitals from which the determinantal functions are constructed. A priori, we should not know them until the optimization has been performed. While a traditional approach — in the mathematical sense — is to approximate molecular orbitals with Legendre, Chebysheb,... polynomials, the chosen path has been guided by the use of chemical intuition. That is, atomic deformations arising from the transferability of an atom from being isolated to an atom in a molecule should be small. Hydrogenoid eigenfunctions suggest us the shape of the molecular orbitals assuming that the atomic character is almost preserved, which is a good approach, based on chemical experience. Any multi-electron system may be described by hydrogen-like orbitals. A clear deficiency of this is that hydrogen functions do not span the whole space of one electron functions. Following this principle we find two historical approaches. Molecular Orbital Theory constructs orthonormal molecular orbitals as a linear combination of atomic orbitals (LCAO). There is no loss of generality on imposing orthonormality to the spinorbitals because what matters is to span the mono-electronic space, then a basis of linearly independent spin-orbitals is the smallest basis that spans the whole mono space. Valence Bond Theory (VB), on the other hand, makes combinations of atomic orbitals always trying to preserve their locality. The purpose is to let the wavefunction mimic Lewis objects to evoke chemical intuition. VB escapes the optional orthonormality condition set for Slater determinants and translates them to the coefficients of determinants.

Both MOs and VB present a path to get to the real wavefunction. If we could expand the wavefunction with an infinite number of atomic orbitals (following the rules of each approach) we would end up with the same result, but we are limited to a finite set of combinations so the chosen approach matters. Sometimes one approach converges faster to the true wavefunction, sometimes the other is more appropriate, sometimes both are equal, and sometimes none is good. Our election has been MO theory for its wide software offer but VB also has a number of interesting traits, like an explicit separation of exchange.

1.5.1 The Linear Combination of Atomic Orbitals (LCAO) expansion

A linear expansion of any molecular orbital in a basis $\{\chi_\nu\}_{\nu=1}^{N_b}$ of atomic orbitals χ_ν with finite cardinality N_b

$$\phi_i = \sum_\nu^{N_b} C_{\nu i}\chi_\nu \tag{1–94}$$

transforms the orthonormalization condition for molecular orbitals into

$$\delta_{ij} = \sum_{\mu\nu} C_{\mu i}S_{\mu\nu}C_{\nu j} = \mathbb{1}_N \quad \equiv \quad C^\dagger SC = \mathbb{1}_N, \tag{1–95}$$

where we can identify the overlap matrix $S_{\mu\nu} = \int \chi_\mu^*(\mathbf{r})\chi_\nu(\mathbf{r})\,d\mathbf{r} \neq \delta_{\mu\nu}$ of atomic orbitals. HF differential equations are continuous and need discretization to be solved algorithmically. Therefore, we search for solutions in a finite dimensional space. The best we can do is to impose the equality $F\psi = \varepsilon\psi$ in the finite dimensional space E^{N_b} and hope that increasing the dimension of the linear space will drive to the solution in the infinite space. Projecting a spinless version of (1–90) into E^{N_b}

$$F\sum_\nu C_{\nu i}\chi_\nu(\mathbf{r}_i) = \varepsilon_i \sum_\nu C_{\nu i}\chi_\nu(\mathbf{r}_i) \tag{1–96}$$

and multiplying with a test function $\chi_\mu^*(\mathbf{r}_i)$

$$\chi_\mu^*(\mathbf{r}_i)F\sum_\nu C_{\nu i}\chi_\nu(\mathbf{r}_i) = \varepsilon_i \sum_\nu C_{\nu i}\chi_\mu^*(\mathbf{r}_i)\chi_\nu(\mathbf{r}_i). \tag{1–97}$$

Forcing the equality to hold after integration yields a weak formulation of the problem for each electron i:

$$\sum_\mu C_{\nu i}\int \chi_\mu^*(\mathbf{r}_i)F\chi_\nu(\mathbf{r}_i) = \varepsilon_i \sum_\nu C_{\nu i}S_{\mu\nu}. \tag{1–98}$$

The Fock operator in this basis ($\langle\chi_\mu|F|\chi_\nu\rangle$), including boundary conditions, becomes

$$\begin{aligned} F_{\mu\nu}(\mathbf{r}_i) &= \frac{1}{2}\int \nabla\chi_\mu^*(\mathbf{r}_i)\cdot\nabla\chi_\nu(\mathbf{r}_i)\,d\mathbf{r}_i - \sum_A \int \chi_\mu^*(\mathbf{r}_i)\frac{Z_A}{\mathbf{r}_i - \mathbf{R}_A}\chi_\nu(\mathbf{r}_i)\,d\mathbf{r}_i \\ &+ \sum_{\sigma,\lambda}^{N_b}\sum_j^N C_{\sigma j}C_{\lambda j}\left[(\mu\sigma|g_{ij}|\nu\lambda) - (\mu\sigma|g_{ij}|\lambda\nu)\right] \\ &= h_{\mu\nu} + \sum_{\sigma,\lambda}^{N_b}\sum_j^N (J_{\mu\nu} - K_{\mu\nu}) \end{aligned} \tag{1–99}$$

after integration by parts of the kinetic integral. The mono-electronic operator looses its diagonality ($h_i \to h_{\mu\nu}$) and bi-electronic integrals [5] are now four-center (μ,ν,σ,λ) integrals, not (i,j) integrals. Eq. (1–98) reads now

$$\sum_\nu F_{\mu\nu}C_{\nu i} = \varepsilon_i \sum_\nu S_{\mu\nu}C_{\nu i}. \tag{1–100}$$

The minimization problem is converted into a generalized eigenvalue problem solving the conditions for first order optimality. That gives us any type of critical point (minima, maxima or saddle) because the energy functional is not necessarily convex. Moreover, the minimum might not be global. All this tells us that somehow we have to integrate in the SCF process that it is a minimization. Only a few (N_b), of the possible, eigenvalues are obtained with each iteration of the N coupled equations ($F(C)C = SCE$) that appear for a system with N electrons. Still they are a greater number than electrons. The aufbau principle states that the lowest eigenvalues $\varepsilon_1 < \cdots < \varepsilon_N$ correspond to an infimum of

$$\mathcal{E}^{HF}(C) = \sum_i^N \sum_{\mu,\nu} h_{\mu\nu}C_{\mu i}C_{\nu i} + \sum_{\sigma,\lambda}^{N_b} \sum_{i,j}^N (J_{\mu\nu} - K_{\mu\nu})C_{\mu i}C_{\nu i} \tag{1–101}$$

with an error quadratic in the error of C. The density matrix γ_1, represented in the basis of atomic orbitals

$$\begin{aligned}\gamma_1(\mathbf{x},\mathbf{x}') &= \sum_i^N \psi_i(\mathbf{x})\psi_i^*(\mathbf{x}') = \sum_i^N \left[\sum_\mu^{N_b} C_{\mu i}\chi_\mu(\mathbf{x})\right]\left[\sum_\nu^{N_b} C_{\nu i}\chi_\nu(\mathbf{x}')\right] \\ &= \sum_\mu^{N_b}\sum_\nu^{N_b}\left[\sum_i^N C_{\mu i}C_{\nu i}\right]\chi_\mu(\mathbf{x})\chi_\nu(\mathbf{x}') = \sum_{\mu,\nu} P_{\mu\nu}\chi_\mu(\mathbf{x})\chi_\nu^*(\mathbf{x}')\end{aligned} \tag{1–102}$$

as the Fock–Dirac matrix $P = CC^\dagger$ is assembled populating only the N molecular orbitals $\psi_1,\cdots,\psi_N$ of lower energy. The remaining $N_b - N$ orbitals are said to be virtual.

Basis sets of polyatomic molecules are divided in subclasses centered around atoms $A, B, \cdots$

$$\left\{\chi_1^A(\mathbf{r}-\mathbf{R}_A),\cdots\chi_{n_A}^A(\mathbf{r}-\mathbf{R}_A),\chi_1^B(\mathbf{r}-\mathbf{R}_B),\cdots\chi_{n_B}^B(\mathbf{r}-\mathbf{R}_B),\cdots\right\} \tag{1–103}$$

to sum a total number $n_A + n_B + \cdots + n_M = N_b$. Among the basic information needed to perform an ab initio calculation stands the set of nuclear coordinates and the atomic basis sets. Whenever the atoms displace integrals have to be recomputed.

Reduction of the HF model to a finite model as above is a completely general procedure, known among the numerical analysis community as a Galerkin

5 Integrals over spatial orbitals $(\mu\sigma|\nu\lambda)$ continue being in physicists' notation.

method, so any basis could be applied. Good basis sets should have the following desirable properties: a small basis set should be enough to properly describe the ground state eigenfunction, and the functions of the basis set should be easy to compute and integrate. There is consensus that the angular part has to be an harmonic function. The radial part, dependent in the potential, generates more dispute.

Hydrogenoid orbitals are very limited to represent electrons immerse in a cloud of electrons. For instance, electrons close to nuclei shield the nuclear-electron attraction felt by outer electrons. Another drawback is that they only form a complete basis when the unbounded eigenfunctions are accounted. And unbounded states do not interest us because we aim to study bounded states.

1.5.2 Slater Type Orbitals (STO)

Slater came up with a set of simple functions (STO)

$$\chi_{\zeta_{nl}lm} = \frac{(2\zeta_{nl})^{\frac{3}{2}}}{\sqrt{(2l+2)!}}(2\zeta_{nl}r)^{l}\exp\left(\frac{-\zeta_{nl}r}{n}\right)Y_{lm}(\theta,\varphi) \tag{1–104}$$

that approximate the molecular orbitals tuning the atomic numbers $\mathbb{N} \ni Z \to \zeta \in \mathbb{R}^{+}$ albeit loosing orthogonality. The combined use of variable quantum numbers n and exponents ζ_{nl} allows to retain high accuracy with small sets. Even with the crudest approximation, that is, using a minimal basis[6], we can obtain quite accurate energies. The main obstacle for their use in large systems is that STO integrals do not admit analytical expressions.

Gaussian Type Orbitals (GTOs)

The tensor $(\sigma\mu|g_{ij}|\nu\lambda)$ is a $6N$ dimensional integral that constitutes the core of any electronic structure calculation. Its scaling is $\mathcal{O}(N_b^4)$. A pivotal moment in the development of electronic structure software occurred when Boys decided to change the basis sets, replacing the exponential with gaussian. Only after that large scale calculations became possible. The advantage of Gaussian functions is the product formula. Four center integrals are reduced to two center integrals and two center integrals are further reduced to one center integrals. All integral evaluations are turned down to computations of one dimensional integrals (Boys functions), which is the bottleneck of all electronic structure calculations, determines to great extent the accuracy achieved in the ground state energy calculation.

6 One basis function per Hartree–Fock orbital ($N_b = N/2$)

1.6 POST–HARTREE–FOCK VARIATIONAL METHODS

1.6.1 Configuration interaction (CI)

We have mentioned, and it is here reinforced, that if $\{\psi_i(\mathbf{x})\}_i^\infty$ forms a complete basis for the mono-electronic space then the set of all Slater determinants filled with electrons in all possible combinations forms an orthonormal basis for the whole space. The most contracted multi-electronic expansion, given by Hartree–Fock theory, condensates N electrons in exactly N spinorbitals. However, correlated electrons cannot be described by a single spinorbital that is *per se* an independent distribution function. Instead, the spread over all possible one-electronic states truly represents any correlated electron. With a number of spinorbitals greater than N, restricted to orthonormalized spinorbitals, the unique way to populate partially spinorbitals is to add more ordered configurations K (determinants) to the total wavefunction:

$$\Psi = \sum_K c_K \Psi_K \tag{1–105}$$

When both single and multi electronic spaces are complete there is no reduction of the original variational space the wavefunction belongs to and we can therefore conclude that diagonalization of H in that basis would be a direct attack to the Schrödinger equation. Down to earth implementations start from a finite number of spinorbitals (Full-CI) to get an exact solution given a basis set, relativistic and Born-Oppenheimer approximations aside. Even so, the scaling is rather acute. Provided N_b spinorbitals have been found previously with Hartree–Fock, the number of partitions (determinants) with N elements of the N_b functions is $N_{det} = \binom{N_b}{N}$. The cost of diagonalization being a cubic power of N_{det}.

It is evident that we have to seek the most important determinants that allow us to converge to the correct solution with a minimum number of them. The single best determinant is the HF determinant and the next are those that interact more with the reference determinant. Only Slater determinants that differ by less than two spinorbitals to the HF determinant should interact directly with it through the Hamiltonian since the later has at most two electron terms. This fact is evident in equation (1–84) and it is explicit in Slater-Condon rules. Not even singly excited determinants interact (directly) with the HF determinant as a consequence of the variational formulation of the HF model (Brillouin theorem). Coupling to other determinants is indirect, rather than being absent. The Hamiltonian encourages a classification of determinants in:

- First-order interacting space: composed of all singly excited determinants where the i-th occupied spinorbital is replaced with a virtual spinorbital a:

$$\Psi_i^a = |\psi_1 \cdots \psi_{i-1} \psi_a \psi_{i+1} \cdots \psi_N\rangle. \tag{1–106}$$

- Second-order interacting space: with all double excited determinants,

$$\Psi_{ij}^{ab} = |\psi_1 \cdots \psi_a \cdots \psi_b \cdots \psi_N\rangle, \tag{1–107}$$

- and so on forth.

The representation of the Hamiltonian in the ordered basis of determinants is a block banded matrix (see table 1).

CISD and CISDT are computational models with functions of the first and second, and up to the third interacting space. A more general CI wavefunction is taken from

$$\mathcal{A}^{CI} = \left\{ \Psi = c\Psi_{HF} + \sum_{i,a} c_i^a \Psi_i^a + \sum_{i,j,a,b} c_{ij,ab} \Psi_{ij}^{ab} + \cdots : c, c_i^a, c_{ij}^{ab}, \ldots \in \mathbb{C}, \langle\Psi|\Psi\rangle = 1 \right\} \tag{1–108}$$

truncating at any level. Fixing the maximum level m of excitations reduces the number of determinants to be the counting of all possible events where we choose to excite $k \leqslant m$ of the N electrons to populate k virtual spinorbitals of the $N_b - N$ pool. If all electrons are excited the result is the same as full CI

$$\binom{N_b}{N} = \sum_{k=0}^{N} \binom{N}{k} \binom{N_b - N}{k}. \tag{1–109}$$

CI minimizes the functional $\mathcal{L}(\mathbf{c};\lambda) = \langle\Psi|H|\Psi\rangle - \lambda\left(\langle\Psi|\Psi\rangle - 1\right)$ with respect to configuration parameters maintaining constant population. We either diagonalize a lower rank CI matrix or solve the system of equations at the critical point

$$\frac{\partial\mathcal{L}(\mathbf{c};\lambda)}{\partial\lambda} = 0, \qquad \frac{\partial\mathcal{L}(\mathbf{c};\lambda)}{\partial c} = 0. \tag{1–110}$$

Excitations involve one, two, and higher electron processes and as we include higher orders the energetic contributions converge monotonically. Inclusion of higher angular momentum orbitals is the same as taking into account the polarization generated by electron correlations. That part is often quoted as *dynamic correlation*. Non-dynamic correlation is present when the HF reference is not good enough to generate all relevant polarizations. To give an example, an state that is degenerate or near to degeneracy (quasi-degenerate states) couples with other states near in the PES to become a unified source of polarizations. These situations are out of the scope of low rank CI approximations. A first step selecting the best references would be necessary for later generating the polarizations with CI. Equilibrium conformations, closed shell, free of degeneracy, frequently achieve convergence with the addition of a few excited configuration spaces, e.g. CISDTQ, that contribute a lot to the energetics.

From a practical perspective convergence is optimal when guess orbitals are the natural orbitals from diagonalizing the first order density matrix.

The cost of cutting the sum over interacting spaces is that the energy does not scale proportionally to the number of electrons. For example, the simulation of

	$\|HF\rangle$	$\|S\rangle$	$\|D\rangle$	$\|T\rangle$	$\|Q\rangle$	$\cdots$	$\|N-3\rangle$	$\|N-2\rangle$	$\|N-1\rangle$	$\|N\rangle$
$\langle HF\|$	H	0	H	0	0	$\cdots$	0	0	0	0
$\langle S\|$	0	H	H	H	0	$\cdots$	0	0	0	0
$\langle D\|$	H	H	H	H	H	$\cdots$	0	0	0	0
$\langle T\|$	0	H	H	H	H	$\cdots$	0	0	0	0
$\langle Q\|$	0	0	H	H	H	$\cdots$	0	0	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\langle N-3\|$	0	0	0	0	0	$\cdots$	H	H	H	0
$\langle N-2\|$	0	0	0	0	0	$\cdots$	H	H	H	H
$\langle N-1\|$	0	0	0	0	0	$\cdots$	H	H	H	H
$\langle N\|$	0	0	0	0	0	$\cdots$	0	H	H	H

Table 1: The matrix representation of H is hermitian. The first matrix element is $\langle HF|H|HF\rangle = E^{HF}$, Brillouin: $\langle S|H|HF\rangle = \langle HF|H|S\rangle = 0$. The upper left rank 3 matrix is the matrix of CISD. Singles do not interact directly with the HF function but indirectly through the double excited functions. $\langle S|H|D\rangle$, $\langle D|H|S\rangle$, $\langle D|H|HF\rangle$,$\langle HF|H|D\rangle$.

a reaction does not treat reactants and transition states on the same footing. The wavefunction is said to be not *size-consistent*. Coupled cluster methods revolve around this problem with a satisfactory solution.

Quite often only a particular spin state interests us. It can be proven with algebraic arguments that all the wavefunctions Ψ_i with different eigenvalue than the desired S

$$S^2\Psi_i = S(S+1)\Psi_i \tag{1–111}$$

will not couple with the true wavefunction Ψ ($\langle\Psi_i|H|\Psi\rangle = 0$). This is true for any operator that commutes with the Hamiltonian. Broadly speaking, determinants need not be eigenfunctions of the squared spin operator S^2. Thus it is not known a priori what spin eigenvalue will have the linear combination of determinants. The task of elucidating which Slater determinants will have correct spin, or non vanishing coupling with the desired spin state function, can be clarified creating adequate linear combinations of determinants: Configuration State Functions (CSFs). CSFs have definite spin eigenvalues. The goal of using CSFs is that spin and spatial symmetry reduces the number of determinants. Another aside is that determinantal implementations can allow crossings of states with different spin eigenvalue S. If one intends to follow a state along a PES the relative position in the list of eigenvectors can change between different computations.

With the CI method we have covered a way to recover accurate wavefunction. However, to achieve a reasonable accuracy, often, a very large number of configurations are needed. The reason is that the virtual orbitals have not been optimized (poorly represented) with the aufbau procedure. Several ways have been came to the rescue selecting the most important configurations, be it with

natural orbitals, perturbation techniques or Multi-Configuration Self Consistent Field (MCSCF). The last one is treated in the following.

1.6.2 Multi Configuration Self Consistent Field (MCSCF)

MCSCF wavefunctions are determinant expansions (1–24) reduced to some fixed rank (K), i.e. the number of determinants forming the wavefunction is K

$$\Psi = \sum_{i_1<\ldots<i_N\leqslant K} C_{i_1,\ldots,i_N}|\psi_{i_1}\cdots\psi_{i_N}\rangle, \tag{1–112}$$

with the distinguishing feature, compared with the CI method, that orbitals where are also optimized. Thus, the set of allowed wavefunctions

$$\mathcal{A}^{\text{MCSCF}} = \Big\{\Psi = c\Psi_0 + \sum_{i,a} c_i^a\Psi_i^a + \sum_{i,j,a,b} c_{ij,ab}\Psi_{ij}^{ab} + \cdots :$$
$$c, c_i^a, c_{ij}^{ab}, \ldots \in \mathbb{C}, \langle\Psi_i^a|\Psi_j^b\rangle = \delta_{ij}^{ab}, \langle\Psi_{ij}^{ab}|\Psi_{kl}^{cd}\rangle = \delta_{ijkl}^{abcd}, \ldots, \langle\Psi|\Psi\rangle = 1\Big\} \tag{1–113}$$

is like $\mathcal{A}^{\text{CI}}$ but each orbital, with origin in an initial HF step for example, is allowed to be optimized keeping the normalization $\langle\psi_i|\psi_j\rangle$. The aim is to obtain better spinorbitals than with HF to allow a faster convergence of the multi-electronic expansion. Remember that HF does not optimize virtual orbitals and that CI expansions do it only one or two orbitals at a time.

The range of selected determinants can go from $K = N$, which is exactly Hartree-Fock, to $K = N_b$, that is Full-CI in that basis. At intermediate values the MCSCF method provides lower upper bounds for the energy than CI truncated at the same level. For practical purposes the MCSCF determinants and orbitals can be fed to a CI calculation of the same rank and both calculations are identical because the orbitals are already optimized and the configuration optimization is supposed to be the same.

The energy has a CI-like expression

$$E = \sum_{J,L} c_J^* c_L H_{JL} \tag{1–114}$$

where $H_{JL} = \langle\Psi_J|H|\Psi_L\rangle$. With the aid of the creation ($a_i^\dagger$) and annihilation (a_i) operators defined by the following operations on the wavefunction

$$a_p^\dagger|\psi_q\rangle = |\psi_p\psi_q\rangle \qquad a_p^\dagger|\psi_p\rangle = 0 \tag{1–115}$$
$$a_p|\psi_p\psi_q\rangle = |\psi_q\rangle \qquad a_q|\psi_p\rangle = 0, \tag{1–116}$$

the Hamiltonian in the spinorbital basis has a neat expression

$$H = \frac{Z_A Z_B}{R_{AB}} + \left[\sum_{p,q} h_{pq}a_p^\dagger a_q + \frac{1}{2}\sum_{p,q,r,s}\langle pr|g_{12}|qs\rangle a_p^\dagger a_r^\dagger a_s a_q\right]. \tag{1–117}$$

When $K = N + 1$ the equations collapse to the Hartree-Fock solution. A Generalized Brillouin Theorem

$$\langle \Psi | [H, a_p^\dagger a_q] | \Psi \rangle = 0 \tag{1–118}$$

dictates that a single orbital modification has no effect because those changes have been accounted by the variational minimization.

An interesting observation is that creation and annihilation operators appearing in the previous expression, when combined to run over all σ spins $E_{pq} = \sum_\sigma a_{p\sigma}^\dagger a_{q\sigma}$, satisfy the commutation relations that are characteristic of the unitary group $U(n)$ generators. E_{pq} united with the spin-adapted generator products

$$\begin{aligned} E_{pqrs} &= \sum_{\sigma \neq \tau} a_{p\sigma}^\dagger a_{r\tau}^\dagger a_{s\tau} a_{q\sigma} \\ &= \sum_{\sigma,\tau} a_{p\sigma}^\dagger a_{q\sigma} a_{r\tau}^\dagger a_{s\tau} - \delta_{qr}\delta_{\sigma\tau} \sum_\sigma a_{p\sigma}^\dagger a_{s\sigma} \\ &= E_{pq}E_{rs} - \delta_{qr}E_{ps} \end{aligned} \tag{1–119}$$

helps us writing the elements of the spin-summed first and second order density matrices

$$\gamma(p|q) = \langle \Psi^{mc} | E_{pq} | \Psi^{mc} \rangle = \sum_{JL} c_J^* c_L \gamma^{JL}(p|q) \tag{1–120}$$

$$\gamma_2(pq|rs) = \langle \Psi^{mc} | E_{pqrs} | \Psi^{mc} \rangle = \sum_{JL} \langle c_J^* c_L \gamma_2^{JL}(pq|rs) \tag{1–121}$$

where $\gamma^{JL}(p|q) = \langle \Psi_J | E_{pq} | \Psi_L \rangle$ and $\gamma_2^{JL}(pq|rs) = \langle \Psi_J | E_{pqrs} | \Psi_L \rangle$ are quoted as the coupling coefficients. They only have -1,0, or 1 values. This makes for an alternative formula of the electronic energy in the spatial-orbital basis

$$\begin{aligned} E &= \sum_{J,L} c_J^* c_L \left[\sum_{p,q}^{K} h_{pq} E_{pq} + \frac{1}{2} \sum_{p,q,r,s}^{K} (pq|g_{12}|rs) E_{pqrs} \right] \\ &= \sum_{p,q}^{K} \sum_{\mu,\nu}^{N_b} \gamma_1(p|q) h_{\mu\nu} C_{\mu j} C_{\nu i} + \frac{1}{2} \sum_{p,q,r,s}^{K} \sum_{\mu,\nu,\sigma,\lambda}^{N_b} \gamma_2(pq|rs)(pq|g_{12}|rs) C_{\mu p} C_{\nu q} C_{\sigma r} C_{\lambda s}. \end{aligned} \tag{1–122}$$

The wavefunction of the ground state is stationary under variations of configuration coefficients and orbitals, under the constraints $\sum_K c_J^2 = 1$ and $\langle \phi_i | \phi_j \rangle = \delta_{ij}$. Optimizing configuration coefficients is the same as in CI, with Lagrange multipliers to conserve normality). On the other hand, the optimal coefficients can be stated as the result of applying a unitary transformation U to the trial orbital coefficients $C = C_{\text{trial}} U$. U is parametrized $U = \exp(R)$ with a skew-symmetric matrix $R^\dagger = -R$. The advantage of this parametrization is that the orbitals are automatically orthogonal, so no constrained optimization is

needed. Also, only $(1/2 N_b(N_b-1))$ entries of the lower triangle of R have to be optimized. It is customary to use a second order optimization with Newton-Raphson. Since the algorithm used to perform this task solves an optimization problem we have to provide a good initial guess.

Practical implementations of MCSCF do not take all the determinants of the expansion (1–112), just a few of them. In principle, MCSCF does not impose systematic rules to guide the selection of the configurations but some approaches have been devised. One of the most successful, Complete Active Space SCF (CASSCF) or Fully-Optimized Reaction Space (FORS), divides the orbital space in two groups: N_v core (or inactive) and N_v valence (or active) orbitals where $N_v + N_c = K$. The resulting wavefunction is of the type [98]

$$\Psi = |\psi_1 \cdots \psi_{N_c}\rangle \wedge \sum_{\{i_1<\ldots<i_{N_c}\}\subset\{N_c+1,\ldots,K\}} C_{i_1\ldots i_{N_v}} |\psi_{i_1} \cdots \psi_{i_{N_v}}\rangle \quad (1\text{–}123)$$

The core electrons are assumed to be well represented by a HF wavefunction whereas the electrons that are more correlated are represented by the set of all configurations that can be created from a pool of active orbitals that are optimized at the same time. It is the same as performing a Full-CI in the space of active orbitals but this time we only diagonalize a subspace of the Hamiltonian matrix (vs CI), only the lowest eigenstates interest us. The virtue of CASSCF resides on the idea of building upon the HF method, the canonical orbital energies can guide us to select properly which orbitals are sufficiently low in energy to be almost unaffected by correlation and which ones are susceptible to be altered by correlation. In any case, chemical intuition is needed to choose correctly the active space, but less than for a general MCSCF calculation. To round up, the energy evaluation and optimization is similar to full MCSCF but now the density matrix is block diagonal with the block corresponding to the core space being an identity matrix and the other block a density matrix on the space of valence orbitals. The same applies for the unitary matrix converting the orbitals between basis.

A final remark: although it is commonly stated that wavefunction methods are fully ab initio they contain some empiricism underlying the choice of the basis functions and active spaces that has relevance because actual applications do not span the whole space so the convergence is conditional on the election.

1.7 DENSITY FUNCTIONAL THEORY (DFT)

A tantalizing idea is to use the electron density instead of the wavefunction. In essence it constitutes a completely different attack to the electronic structure problem, reducing the dimensionality from $(\mathbb{R}^{3N} \times \{+,-\})$ to only $\mathbb{R}^3$ at the expense of not knowing the exact expression for the energy functional in terms of the density, in contrast with wavefunction methods where it is totally known. Simultaneously the wavefunction and the Hamiltonian are approximated.

An informal argument that justifies this theory (for Coulomb potentials) is as follows: if we knew the electron density we would know that nuclei are located at the cusps and by a fundamental result of T. Kato we also would know the charge of the nuclei at those positions. Moreover, the number of particles is known from the density by integration. With that knowledge we could solve the Schrödinger equation for the full space.

Originated as an approximation in the theory of solids the first attempts to build a density functional date back to the decade of 1920, from the early works of Thomas, Fermi and Slater. Only after the seminal paper of Kohn-Sham it becomes evident that the density could have a far reaching application. In the forthcoming years, especially after the adoption in widely used molecular simulation codes, it has been broadcasted to all the disciplines involved in the simulation of multiple electronic systems. Today, it is the preferred method of choice for ab initio simulation of large molecular systems.

1.7.1 Pre–DFT: Thomas–Fermi–Dirac, Slater Xα

Thomas-Fermi Theory comprises various models that try to find the density and ground state energy of a system with a large number of electrons $N \to \infty$. Augmenting at the same time the positive charge bath $Z \to \infty$ entails an exact theory. The central problem is to find a universal functional of the electronic density that is independent of the physical nature of the system. The initial model considered is a gas of electrons in a fixed volume with classical interactions. On that ground the unique energetic terms are kinetic

$$T[\rho] = C_T \int \rho^{5/3}(\mathbf{r})\, d\mathbf{r} \tag{1–124}$$

and Coulomb attraction-repulsion of charges

$$E_C[\rho] = \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\, d\mathbf{r}_1\, d\mathbf{r}_2 + \int \rho(\mathbf{r}) V_{ne}\, d\mathbf{r}. \tag{1–125}$$

Further improvements approximate exchange in an equivalent fashion as Hartree-Fock:

$$E_X = -\frac{1}{2}\int \frac{|\gamma_1(\mathbf{x};\mathbf{x}')|^2}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{x}\, d\mathbf{x}' = -C_X \int \rho^{4/3}(\mathbf{r})\, d\mathbf{r} \tag{1–126}$$

A dimensional justification takes the eigenfunctions of the kinetic energy operator in a large box of volume V

$$\psi_{i,\mathbf{k}}(\mathbf{x}) = \frac{1}{\sqrt{V}} e^{2\pi \mathbf{k}\mathbf{r}} s_i(\sigma) \qquad \mathbf{k} = \frac{\mathbf{n}}{|V|^{1/3}} \quad \mathbf{n} \in \mathbb{N}^3 \tag{1–127}$$

to the limit of a large number of particles maintaining constant density $\rho = \frac{N}{|V|}$. In that setting $E_X = -C_X \rho^{4/3}|V|$

$$\int \rho(\mathbf{r})\, d\mathbf{r}\,[=]\,\frac{V}{V}\,[=]\text{dimensionless}$$
$$\int \rho^{4/3}(\mathbf{r})\, d\mathbf{r}[=]\left(\frac{1}{V}\right)^{4/3} V = V^{-1/3} = L^{-1}[=]\int \frac{|\gamma_1(\mathbf{x};\mathbf{x}')|^2}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{x}\, d\mathbf{x}'$$
$$\int \rho^{5/3}(\mathbf{r})\, d\mathbf{r}[=]\left(\frac{1}{V}\right)^{5/3} V = A^{-1}[=]\int |\nabla_{\mathbf{r}}\rho(\mathbf{r})|^2\, d\mathbf{r} \tag{1–128}$$

Thus far the energy functional achieves an almost correct form. Later corrections were added to the kinetic energy.

1.7.2 Hohenberg-Kohn density functional

Even if previous work had been done using the density as the fundamental physical variable, the modern theory of DFT is thought to be born only after the groundbreaking results of Walter Kohn and P. Hohenberg (HK) in 1964 [82]. They established a bijective mapping between the density and the external potential $v : V_{ne} = \sum_k^N v(\mathbf{r}_k)$ first, and second they defined a functional of the density that fulfills the variational principle — under restricted conditions—.

The first fact we have to remark is that the only term that prevents us from knowing the total energy of a *universal*[7] N electron system is the potential $v(\mathbf{r}_k)$ created by the nuclei upon an electron placed at $\mathbf{r}_k$ and a constant V_{nn} we shall omit altogether for a more clear exposition. The rest of the energy is intrinsic to the electronic system. This way

$$E = E_{int} + E_{ext} = E_{int} + \langle\Psi|V_{ne}|\Psi\rangle. \tag{1–129}$$

In a precedent theorem of Rosina it was already recognized that the energy for any unknown two-particle Hamiltonian is known from the second order density matrix alone, or the tuple of density, 1RDM and 2RDF. Although the one particle density is not enough to reconstruct the Hamiltonian.

Under variations δv, the Hamiltonian is modified just by $\delta H = \delta V_{en}$. Variations of the ground state energy are reduced to

$$\delta E = \int \frac{\delta E}{\delta v(\mathbf{r})}\delta v(\mathbf{r})\, d\mathbf{r} = \langle\delta\Psi|H-E|\Psi\rangle + \langle\Psi|\delta V_{en}|\Psi\rangle + \langle\Psi|H-E|\delta\Psi\rangle$$
$$= \langle\Psi|\delta V_{en}|\Psi\rangle = \int \rho(\mathbf{r})\delta v(\mathbf{r})\, d\mathbf{r}. \tag{1–130}$$

The last equality can be restated as: the density is the functional derivative of the energy with respect to the external potential when the state is non-degenerate. The proof for the converse relation is proved by *reductio ad absurdum* supposing that two ground states with respective ground state energies

7 In the sense that any N electron system has the same energy functional

and Hamiltonians, that obviously differ only in δV_{en}, share the same density. Relying on the variational principle the two densities have to be different if the wavefunctions that minimize the Hamiltonians are non-degenerate, violating our initial assumption. In conclusion the claim

$$v(\mathbf{r}) \iff \rho(\mathbf{r}) = N \sum_{\sigma_1} \int |\Psi|^2 \, d\mathbf{x}_{i>1} \tag{1–131}$$

follows.

In our treatment of the N electronic problem we can start from the density or from the potential, in fact the second seems more affordable. Despite of this there is no clue of how we can select a potential that bounds the N electrons. In the following HK go around the problem focusing on the density functional drawing on $\rho(\mathbf{r}) \to v(\mathbf{r}) \to \mathcal{H}_{el} \to \Psi$. In a deliberate act of optimism we assume that the energy $E(v)$ is known for a family of potentials that form a linear space. Hiding the complexity of computing the energy terms that depend on objects more complicated than the density and keeping the external potential fixed, which requires only knowledge of the electronic density

$$E_{int} = F_{HK}[\rho] = E[v] - \int \rho(\mathbf{r}) v(\mathbf{r}) \, d\mathbf{r}. \tag{1–132}$$

The total energy is formulated minimizing over proper densities

$$E[v] = \inf_{\rho} \left[\int v\rho \, d\mathbf{r} + F_{HK}[\rho] \right]. \tag{1–133}$$

The problem with the HK functional is that the definition of the density functional is vague. The density is thought to be resulting from a wavefunction that is a ground state of a potential v that have a bounded non-degenerate state. Neither the potential nor the densities are contained in a well defined domain. That is not enough clear to base a rigorous theory on it but it is so subtle that no physical effect is found.

1.7.3 Levy–Lieb constrained search

In the Hohenberg-Kohn theorems, one important assumption which has not been mentioned is that, during the minimization, as we vary the density, we assume that it remains v–representable. A v–representable density is a ground state density that can be associated with a Hamiltonian that has an external potential, $v(\mathbf{r})$. It is, however, not clear that an arbitrary density, which integrates to an integer number of electrons N, would be the ground state of a smooth external potential. In order to overcome the v–representability problem, a more general variational routine was proposed independently by Levy and Lieb. Their extended minimization algorithm requires only N-representability of the densities. An N–representable density is one which can be derived from

an antisymmetric wavefunction. This condition, being much weaker than v-representability, was proven to be easily satisfied by an arbitrary density.

Independently, Levy and Lieb proposed a constrained–search formulation that avoids the HK restriction of non degenerate states and v–representatibility constructing the density from a set of normalized wavefunctions that have density ρ that integrates to N. In particular, Lieb recognized the Legendre transform as the appropriate tool to carry this task as a two way optimization. First the Legendre transformation of $E[v]$

$$F[\rho] = \inf_{v}\left[E[v] - \int \rho(\mathbf{r})v(\mathbf{r})\,d\mathbf{r}\right] = \inf_{\Psi\to\Psi_{\rho_0}} E_{\text{int}} = \inf_{\Psi\to\Psi_{\rho_0}} \langle\Psi|T + V_{ee}|\Psi\rangle \quad (1\text{–}134)$$

and secondly the inverse transformation for the total energy

$$E[v] = \inf_{\rho}\left[\int \rho(\mathbf{r})v(\mathbf{r})\,d\mathbf{r} + F[\rho]\right]. \quad (1\text{–}135)$$

The bijective mapping between the potential and the density is interpreted in the context of the Legendre transform as conjugate variables. DFT is a theory of local potentials, both are of Coulomb type, hence the conjugate density is local. This is a point in contrast with Hartree-Fock that has a non-local exchange potential.

1.7.4 Kohn–Sham equations

The idea of Kohn and Sham is to use as a first approximation for the kinetic functional an ideal system of non interacting particles. The non interacting system can then be connected adiabatically to the real one by a parameter $0 \leqslant \lambda \leqslant 1$ in the Hamiltonian

$$H_\lambda = T + \lambda[V_{ne} + V_{ee}]. \quad (1\text{–}136)$$

As long as the same density and total energy are maintained when applying the variational principle to different λ. The real energy is approached when the parameter $\lambda \to 1$.

Whereas the solution of the non-interacting Hamiltonian

$$F_0[\rho] = \inf_{\Psi}\langle\Psi|H_0|\Psi\rangle \quad (1\text{–}137)$$

is not necessarily a monodeterminantal wavefunction, the ensemble of all representable first order density matrices that define any possible outcome of the kinetic operator H_0 are well known (Gilbert). After diagonalization of the density matrices the result is expressed in terms of an infinite linear combination of spin-orbitals

$$F_0[\rho] = \inf_{\gamma_1} \mathrm{Tr}(H_0\gamma_1) = \inf_{\gamma_1}\left[\frac{1}{2}\sum_{i}^{\infty} n_i \int |\nabla\psi_i|^2\right] \quad (1\text{–}138)$$

such that the density is $\rho = \sum_i^\infty n_i |\psi_i|^2$.

In stationary conditions the Euler-Lagrange equations stemming from the energy functional $E(\nu)$ are

$$\delta\widehat{\rho}\left[\frac{\delta\widehat{F}_0}{\delta\widehat{\rho}} + \nu - \epsilon\right] = 0 = \left[-\frac{1}{2}n_i\nabla^2 + \nu - \epsilon_i\right]\psi_i. \tag{1–139}$$

If the density is stationary so is γ_1. The point is that the solution is known exactly. The philosophy behind the Kohn-Sham formalism is to take this as a reference and include the effect of interactions later always trying to get as close to the result as possible with information of the density whereas unknown terms are separated. This way we fence our limited capability to treat the electronic system based only on the density. Increments of λ induce modifications of the first order density matrix, translated to errors in our estimation of the kinetic energy. Also interactions cannot be simulated in their full extension, $\int\int g_{ij}\rho_2 = \int\int g_{ij}\rho\rho - \int\int g_{ij}\rho_{xc}$, only Coulomb interaction of independent particle densities (self-repulsion included)

$$E_H[\rho] = \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,d\mathbf{r}_1\,d\mathbf{r}_2 = \frac{1}{2}\int \rho(\mathbf{r}_1)\nu_H(\mathbf{r}_1)\,d\mathbf{r}_1 \tag{1–140}$$

is accessible. In the DFT community ν_H is known as the Hartree potential. Exchange and correlation, jointly with self-repulsion and kinetic corrections are part of a grab-bag called exchange-correlation energy functional,

$$\begin{aligned} W_{xc}[\rho] &= F[\rho] - F_0[\rho] - E_H[\rho] = (T[\Psi] - H_0[\rho]) + \left(E_{ee}[\rho] - \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,d\mathbf{r}_1\,d\mathbf{r}_2\right) \\ &= T_c[\rho] + \frac{1}{2}\int \frac{\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}\,d\mathbf{r}_1\,d\mathbf{r}_2. \end{aligned} \tag{1–141}$$

Another way to separate the terms of W_{xc}, on the basis that it is a Slater determinant ($\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = \rho_x(\mathbf{r}_1, \mathbf{r}_2)$), is to take the exchange in one side $W_X[\rho] = E_{ee}[\rho] - E_H[\rho]$. The correlation is $W_C[\rho] = T_c[\rho] + (E_{ee}[\Psi] - E_{ee}[\rho])$. Now, adding all terms $F[\rho] = F_0[\rho] + E_H[\rho] + W_{xc}[\rho]$.

The expression for the energy

$$E[\nu] = \inf\left[\int \rho(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + H_0[\rho] + E_H[\rho] + W_{XC}[\rho]\right] \tag{1–142}$$

The variational Euler–Lagrange equation for the KS real system is

$$\delta\rho\left[\frac{\delta F_0[\rho]}{\delta\rho(\mathbf{r})} + \nu_{eff}(\mathbf{r}) - \epsilon\right] = 0 \tag{1–143}$$

where

$$\nu_{eff}(\mathbf{r}) = \nu(\mathbf{r}) + \nu_H(\mathbf{r}) + \frac{\delta E_{XC}[\rho]}{\delta\rho(\mathbf{r})}. \tag{1–144}$$

The external potential is computed solving Poisson's equation from the density. v_{eff} is a local effective potential (multiplicative). The effective potential is chosen to make the density exact. Implementations of KS–DFT approximate the value of v_{eff} without needing to get the truly wavefunction.

$$\left(-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r})\right)\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{1–145}$$

The process of solving the Kohn–Sham equations is self–consistent. KS orbitals ψ_i are a completely different thing from HF orbitals. They are mono-electronic functions of a non-interacting system associated with the real one. Recently they were proved to yield good optical transition energies.

All in all DFT comprises a theory that is in principle exact. Only the exchange correlation functional escapes our knowledge. The approach to deduce its analytical expression is the major concern of its developers.

1.7.5 Functionals

Exchange-correlation functionals are approximated with varying levels of sophistication. The most primitive functionals, inherited from the praised Thomas-Fermi models, will be local. After that the non-locality of the HF exchange is increasingly included. Usually more evolved expressions perform better but the improvement is not guaranteed as we evolve in the ladder. This fault is indebted to the non existence of a variational principle for approximated functionals.

Much of the development of functionals has been guided more by numerical purposes than from mathematical verification. Some functionals do not respect the N-representability ending in the loss of variationality.

Local Density Approximation (LDA)

Already in their paper HK used a LDA for a non-uniform electron gas

$$W_{XC} = -\int \rho(\mathbf{r})\varepsilon_{XC}(\rho(\mathbf{r}))\,\mathrm{d}\mathbf{r} \tag{1–146}$$

based on the formula for a uniform electron gas i.e. a system of interacting particles with constant density, with energy per particle $\varepsilon_{XC}(\rho(\mathbf{r}))$. Comparing with the previous equation for the TF model the exchange energy per particle is $C_X\rho(\mathbf{r})^{1/3}$. There is no expression for the correlation energy as for the exchange but formulas exist (VWN,PW92) from interpolating Monte Carlo simulations of UEG.

When incorporated in the KS recipe: (i)the density of a non-uniform non-interacting gas is computed with $\lambda = 0$, then (ii) interactions, based on a uniform gas, are applied to the inhomogeneus density.

LDA is good for slowly varying densities, riding out molecules where sharp peak densities appear and high accuracy is needed. In solids the results are

better but not because the density is more smooth, which is false. Cancellation errors between exchange and correlation are the reason for its success. While around 80% of the exchange correlation is exchange and 20% is correlation, correlation is underestimated and exchange overestimated in a smaller percentage but at the end nearly cancels.

Generalized Gradient Approximations

Attempts to formalize the nonlocality of the interaction began pointing to a power series expansion of $\rho(\mathbf{r})$ starting with the gradient and then going to higher order terms. The original Generalized Expansion Approximation was

$$W_{xc} = W_{xc}^{LDA} + \int C_{XC}(\rho(\mathbf{r}))\frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})^{4/3}}. \tag{1–147}$$

An aside effect that goes in detriment of the accuracy, it can be even worse than LDA, is that the sum rules of the exchange hole (negativity and integration to -1) and correlation hole (integration to 0) are not always met. Then, a non trivial discovery was that the series expansion had to be left apart to look for more diverse functions of the density, and gradient, that are consistent with the fundamental properties of the correlation holes. Generalized Gradient Approximations (GGA) come in many flavors but all share a similar expression

$$W_x = \int \rho(\mathbf{r})\varepsilon_x(\rho(\mathbf{r}), s(\mathbf{r})), \qquad s(\mathbf{r}) = \frac{|\nabla\rho(\mathbf{r})|}{2\rho(\mathbf{r})k_F} \tag{1–148}$$

with dependence on a dimensionless gradient $s(\mathbf{r})$, where $k_F = (3\pi^2\rho(\mathbf{r}))^{1/3}$ is the Fermi momentum. GGA reflects how important is to model holes correctly.

Approximate functionals of this type, PBE, PW91, or B88, each one of them tries to satisfy a different set of exact conditions when densities are smooth and not very small. In those situations GGAs give reliable results for strong bonds, reducing the overbinding tendency of LDA, but for van der Waals interactions fail. Nor have GGAs achieved the claimed accuracy, of around 1 kcal/mol, needed to discern processes of chemical nature.

Hybrid functionals

Hybrid functional approximations came to birth from the observation [17] realized by Becke that while LDA functionals exaggerate the delocalization of the exchange hole, with correlation localizing the hole progressively as the coupling parameter approaches $\lambda = 1$, GGA localizes it excessively at any value of λ. Therefore, a fraction of the exact hole $W_x^{HF}[\{\psi\}]$ energy from a Hartree-Fock calculation is used for not incurring in the error.

$$W_x[\rho] = aW_x^{HF}[\{\psi\}] + bW_x^{GGA}[\rho] + (1-a-b)W_x^{LDA}[\rho] + cW_c^{GGA}[\rho] + (1-c)W_c^{LDA}[\rho] \tag{1–149}$$

An inflection point happened with the hybrid functionals in the way parameters were fitted. Until their appearance only atomic ionization energies were part of the fitting but from then on molecular data was also responsible for the adjustment of the optimal parameters.

Range-separated and dispersion-corrected functionals

The corrections included by a series expansion revolve around the vicinity of a point $\mathbf{r}_1$ where the integration is done but truly non-local effects like charge transfer or dispersion require information of the density at distant locations. Often, DFT over-estimates and over-stabilizes charge transfer. The situation with dispersion interactions is more dramatic, it has resisted the assault of local or semi-local density functional approximations.

A limit that is very relevant for the development of functionals that correctly describe long-range interactions is the dissociation limit (Heitler London). The well known $1/R^6$ fall off is not respected by those functionals, instead they have an exponential decay. A strategy based on a split Coulomb operator, in the same vein as the Ewald summation,

$$\frac{1}{r_{12}} = \frac{\operatorname{erfc}(\omega r_{12})}{r_{12}} + \frac{\operatorname{erf}(\omega r_{12})}{r_{12}} \tag{1–150}$$

provides a good starting point. The philosophy behind this separation is that local or semi-local DFT approximations correctly treat the short-range electron correlation, while the long-range correlation is better treated by wavefunction methods.

The instantaneous electron-electron interactions distort the electron cloud charge creating temporal (electronic excited states) atomic formal charges that interact with other elements creating an attraction depending on the nature of the charges. It might seem a good option to put an eye on corrections to the exchange-hole, after all GGA functionals have benefited substantially from this approach. Expanding the pairwise interaction of exchange-holes with a perturbation expansion. This yields interaction of multipoles that can be truncated to second order (dipole-dipole interactions). A parameter-free functional, XDM [19, 139].

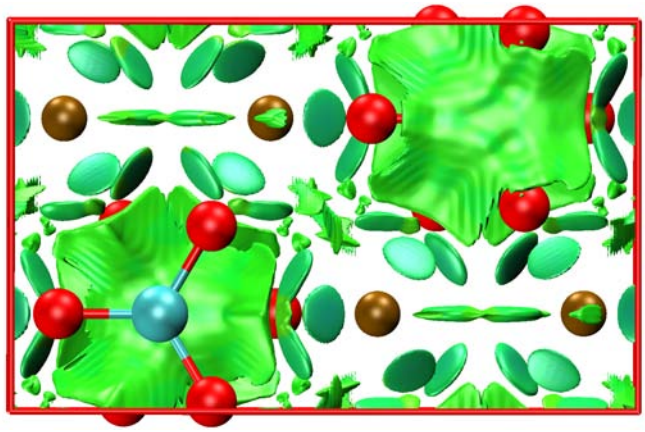
The quest for more accurate functionals has not ceased, and new functionals have continued appearing at a high pace. Among the functionals that we do not treat but deserve a mention are meta-GGA, pure exchange (EXX-based), and double-hybrid functionals. All these functionals can be organized in a hierarchy 2 that slowly approximates to the definitive universal functional, idealized as Jacob's ladder to heaven in the mind of Perdew. When confronted to demanding systems, DFT approximations have sometimes kicked over the traces, challenging the established philosophy behind DFT The fitting of parameters against large datasets of molecules, the use of exact exchange or the theme of debates today but should not hide from us the fruitful understanding that we have gained in the process.

Table 2: Functional Jacob's ladder. s is the reduced density gradient and τ is the kinetic energy density. EXX-based are exact exchange functionals. Hyper-GGA do not resemble GGA at all, the functional exchange is completely replaced by the exact exchange and the non-locality is moved to the correlation functional.

Rung	Name	ε_{XC}
1	L(S)DA	$\varepsilon_{XC}(\rho)$
2	GGA	$\varepsilon_{XC}(\rho, s)$
3	meta–GGA	$\varepsilon_{XC}(\rho, s, \tau)$
4	Hybrid–GGA	$\varepsilon_{XC}(\rho, s, \tau) + E_X^{HF}$
4	Hyper–GGA, EXX-based	$\varepsilon_C(\rho, s, \tau) + E_X^{HF}$
5	Double Hybrid, RPA	$\varepsilon_{XC}(\rho, s, \tau) + E_X^{HF}$ + unoccupied orbitals
-	Non-local (curvature), GLDA1	$\varepsilon_{XC}(\rho, \eta)$

2 SOLID STATE CHEMISTRY

CONTENTS



Given the type and location of atomic elements in the structure of a material, the electronic density and all the relevant properties of the ground state can be determined[a].

a Non relativistic, no external field, and BO approximations

2.1 MATERIALS RESEARCH

The history of our evolution is tied to our ever increasing use of materials. Today our lives are invaded by new smartphones, computers, ...but what is not stressed enough is that all that development has been seeded by research on materials at a nanoscale level.

Some minerals have nearly perfect polyhedric shape that induces us to think that there is a repeating pattern at the atomic level. Why would atoms distribute periodically in their most stable configurations. That raises a complicated question, difficult to solve even in the most naive models. The number of particles to simulate is in the order of Avogadro's number.

Symmetry takes an important role simplifying the problem. It was Kepler, 400 years ago, who first observed that symmetry underlies the disposition of atoms or molecules in condensed matter while inspecting snowflake patterns. We know today that symmetry is present everywhere. Many solids in nature are periodic arrangements of atoms presenting a type of space group symmetry [75].

2.2 IDEAL CRYSTALS

Perfect monoatomic crystals are periodical arrangements of atoms in a lattice defined by the positions

$$\mathbf{R} = \sum_{i=1}^{3} n_i \mathbf{a}_i \qquad \forall n_1, n_2, n_3 \in \mathbb{N} \tag{2–1}$$

that can be generated by translation in any of the three dimensions with the basis vector $\mathbf{a}_i$ a number of times n_i. The region enclosed by the vectors has enough information to regenerate the crystal ad infinitum by repetition. Any point $\mathbf{r}$ of the crystal is reduced to an interior point of the primitive cell $\mathbf{r}_p$ by translation: $\mathbf{r} = \sum_{i=1}^{3} n_i \mathbf{a}_i + \mathbf{r}_p$. When the atoms are not placed at the lattice points the coordinates can be specified by a set of fractional coordinates x, y, z such that

$$\mathbf{r}_p = x\mathbf{a}_1 + y\mathbf{a}_2 + z\mathbf{a}_3 \qquad \forall x, y, z \in [0, 1). \tag{2–2}$$

If the volume V_Ω of cell is minimal we call it a *primitive cell*.

A crystal can have more symmetries than translation. The additional symmetry operations form groups with point symmetry, at least one point of space is left fixed. Those point group operations $R_i\Psi(\mathbf{r}) = \Psi(\mathbf{r}')$ have a matrix representation A that transforms linearly the coordinates $\mathbf{r}$ of one atom to another position $\mathbf{r}' = A\mathbf{r}$. Only a select group of rotations, reflections or combinations of those operations with translations are compatible with an infinite periodicity. The operations defined by Wigner-Seitz symmetry operators

$$\{R_i, T_{\mathbf{R}}\}\Psi(\mathbf{r}) = \Psi(A\mathbf{r} + \mathbf{R}) \tag{2–3}$$

rotate and translate with a vector of the lattice the coordinates of the atoms.

Primitive cells are not unique, however all of them contain only a single point of the Bravais lattice (or atom). There is a unique primitive cell (Wigner-Seitz cell) created via Voronoi partitioning that that satisfies the desirable property of being invariant under the action of all symmetry operations that can be possible in the crystal. Other less compact choices of cell are also common to have a more clear representation of the crystal structure.

In three dimensions fourteen Bravais classes of lattices arise. Adding a motif to the lattice points can increase the number of point groups to 32. When they are combined the set of all possibilities is 230 space-group types. The space-group is not totally characterized until we set the cell parameters: moduli and angles of the primitive vectors.

The dual of the Bravais lattice has many applications, among them the understanding of crystal diffraction and electron density functions with the periodicity of a Bravais lattice. The reciprocal lattice is in itself a Bravais lattice

$$\mathbf{G} = \sum_{i=1}^{3} k_i \mathbf{b}_i \qquad \forall k_i \in \mathbb{N} \tag{2–4}$$

with primitive reciprocal vectors such that $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$. An analogue of the Wigner-Seitz cell in the reciprocal lattice is the first Brillouin zone of volume $V_{\Omega^*} = \frac{8\pi^3}{V_\Omega}$.

Sooner or later one has to admit that the crystalline order is an idealization of the arrangement that atoms exhibit in macroscopic solids. Even under the assumption that the most stable structure is achieved when atoms have a perfect order, there is no crystal free of defects at finite temperature. Solely the entropic contribution of lattice vibrations can provoke a transition to another nuclei disposition. The often mild pressure conditions can turn drastic in some places and favor exotic structures.

2.3 EXPERIMENTAL STRUCTURE DETERMINATION: CRYSTALLOGRAPHY

2.4 DIFFRACTION

X-ray radiation is of particular interest for its similitude in wavelength to interatomic spaces that allows us to "see" the atoms. The usefulness of X-ray techniques relies on a proper description of rays-matter interactions. When a X-ray beam strikes a crystal, the beam is absorbed or scattered from the crystal. Atomic nuclei are thousands of times smaller than the diameter of the atoms and they mostly do not participate in the scattering process. Scattered photons may or may not undergo an energy transfer, both are termed inelastic and elastic scattering respectively. If the wavelength of the incident beam and the

lattice spacing are comparable constructive interference occurs. The process is called diffraction.

The path difference is conditioned by atomic spacings and the orientation of the beam. W. H. Bragg and his son used a simple model of atoms contained in equally spaced planes and found a law,

$$\lambda = 2d \sin\theta, \tag{2–5}$$

that rules the conditions needed to be satisfied for diffraction to happen. The wavelength of the beam (λ) must equal twice the interplanar distance (d) multiplied by the sine of the incident beam angle (θ). The same law works as well for neutron or electron diffraction.

2.5 ELASTIC X–RAY DIFFRACTION

There are two main techniques associated to X–Ray Diffraction. One is single crystal diffraction and the other is powder diffraction. A powder difractogram has not enough peaks to correctly determine the huge number of density modeling parameters. Our data has been gathered with the first technique.

The orientation of scattered X-Ray beams is defined by the scattering vector $\mathbf{H} = 2\pi\mathbf{G}$. Scattered peaks are collected by a bi-dimensional CCD detector. The amplitude, $A(\mathbf{H})$, of the diffraction peaks is proportional to the Fourier transform of the thermally averaged electron density, $\rho^{dyn}(\mathbf{r})$, in the unit cell:

$$I(\mathbf{H}) = |A(\mathbf{H})|^2 \approx |F(\mathbf{H})|^2 = \left| \int_V \rho^{dyn}(\mathbf{r}) e^{2\pi i \mathbf{H}\cdot\mathbf{r}} \, d\mathbf{r} \right|^2 . \tag{2–6}$$

The first step towards achieving a high quality density description is obtaining a good crystal. Another less relevant factor is the type of crystal (ionic crystals usually diffract better than molecular ones with the same quality, because they have tightly bounded electrons) and mosaicity.

Obviously we cannot realize an infinite number of measurements. The limit

$$|\mathbf{H}|_{max} = \frac{2\sin\theta_{max}}{\lambda} \tag{2–7}$$

is set by the type of radiation used and the geometry of diffractometers. The signal is each time weaker as we augment the incidence angle. As a result, we have to truncate the Fourier sum. A huge number of peaks need to be measured to complete the Fourier sum as much as possible. Radiations with small wavelengths increase the number of measurable peaks but the intensity is inversely proportional to the wavelength, so the expected improvement is canceled. The decreasing ratio of amplitudes is emphasized with the temperature.

The main problem comes here. The structure factors are complex numbers of which we only have information about its complex product, $|\mathbf{F}|^2$. The phase has been lost during the experiment and we need to recover it anyway to solve

the structure and density properly. The relevance of the phase information is crucial because most of the information is coded in the phase.

Once the diffraction peaks are obtained, absorption and other experimental effects introduced in the intensities are corrected.

The most valuable experimental technique for the determination of the 3D structure of crystalline samples is X-ray diffraction (XRD). It is of utmost importance to allow the discovery of new materials for the next technological era, drugs to treat human diseases or investigate minerals underpinning the geology of Earth. Nowadays X-ray crystallography is recognized as an essential tool.

2.6 DENSITY MODELING

The unit cell dynamical density can be formally decomposed in pseudo-atomic densities and Dirac's deltas at average atomic positions. The dual of pseudo atomic densities are atomic form factors. Atomic form factors include the Debye–Waller factor, that measures the average thermal motion. Atomic form factors are diminished with increasing Bragg angles. Even at very low temperatures. Heavy atoms have a longer tail, although the maximum angle achieved is barely increased.

Thermal factors are deconvoluted from the static density using a harmonic model of nuclei motion. Low temperature is needed to decrease nuclei motion. Thus the harmonic approximation becomes a better model and the deconvolution of thermal factors from static density fitting parameters is improved. The static density has relativistic, matrix and correlation effects included, and does not correspond to the ground state. The static density is described by a superposition of pseudo-atomic densities

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_{at}} \rho_{at}(\mathbf{r}). \tag{2–8}$$

There are three basic models to refine the density with different levels of sophistication: A naïve approach, the IAM model, is to consider spherical atomic densities. This approximation is good for core electrons. The difficult part to model is the valence density. The aim of other models is to recover the anisotropy of the valence density. In our next approximation, the density is allowed to contract or to expand, that is the *kappa* model. If we go further, we can allow the density to be non spherical adding multipolar expansions, the multipolar model.

The Maximum Entropy Method (MEM) is an alternative to the multipolar model to refine charge densities. It treats directly the dynamical density. The dynamical density obtained can be used [131] to apply QTAIM analysis but we should remember that it is not the same as the static density.

The static density is refined following a least-squares minimization of

$$M = w_{\mathbf{H}} \left(F_{obs}(\mathbf{H})^2 - F_{theo}(\mathbf{H})^2 \right)^2 \tag{2–9}$$

the difference between observed and calculated intensities, each one has a weight $w_{\mathbf{H}}$.

2.6.1 The Independent Atom Model (IAM)

Under the isolated spherical approximation (**IAM**) the density of each atom is build up from the radial part of *ab initio* wave functions . Therefore, the atomic scattering factors are known from Hartree–Fock calculations of ground state free atoms. Even though this model is simplistic it works well for heavy atoms owing to the small proportion of valence shell electrons against total density. For hydrogen atoms the approximation is not good enough even to describe properly atomic coordinates. They have only one electron and their contribution is usually too small. **IAM** overestimates their bonding because the unique electron that it has is displaced towards the bond and therefore the centroid that determines the atomic position is shifted in the same direction. The problem remains even with better models. When the system has hydrogen atoms, the X-ray parameters refinement has to be complemented with neutron diffraction data.

Only thermal parameters and atomic coordinates are refined within this model.

The envelope of core electron form factors do not decrease much with Bragg's angle incrementation. The easiness of the **IAM** model suggests to perform first a refinement with high angle data and then proceed with low angle data to refine the valence shell charge. The problem of this *high order refinement* is that the intensity of peaks is very small.

2.6.2 Kappa formalism

Two new parameters have to be refined in the *kappa* model. The κ parameter allows the density to contract or expand scaling the radial function. $\kappa^3 \rho_{val}(\kappa r)$ is normalized to be a one electron density.

$$\rho_{at}(\mathbf{r}) = \rho_{core}(r) + P_{val} \kappa^3 \rho_{val}(\kappa r) \tag{2–10}$$

P_{val} is the valence shell population parameter. It contains information to analyze charge transfer.

The density is spherically symmetric.

2.6.3 Multipolar pseudo-atom model

The kappa model does not recover angular anisotropy at all. Atoms in a crystal are surrounded by others that disturb the electronic density of its

neighbors. Some directions are preferred for interatomic interaction. Nearest neighbors of different electronegativity induce charge transfer and dipole moments appear. Non spherical orbitals produce the same effect. Is there any function that fills the gap between the theoretical and experimental densities $\Delta\rho(r) = \rho(r) - \rho_0(r)$? The framework [181] to generalize the atomic scattering factors was set by Stewart, after some previous work with multipolar functions. Years later Hansen and Coppens developed the definitive model to extract accurate electronic densities [38, 76]. There are three implementations with similar characteristics: `MOLLY`, `LSEXP` and `POP`.

The pseudo-atomic multipolar density is

$$\rho_{at}(\mathbf{r}) = \rho_{core}(r) + P_{val}\kappa^3\rho_{val}(\kappa r) + \sum_{l=0}^{l_{max}} \sum_{m=-l}^{m=l} \kappa'^3 R_{nl}(\kappa' r) P_{lm} Y_l^m(\theta, \phi). \quad (2\text{–}11)$$

It is a fuzzy atomic partitioning

The first two terms are the same found in the *kappa* model. They contain all the charge of the pseudo-atoms. The last term is a multipolar expansion that interchanges charges between the lobes of real spherical harmonic functions $Y_l^m(\theta, \phi)$.

The density fitting is very sensitive to the choice of the optimal radial function, $R_{nl}(\kappa' r)$. The monopole P_{00} is usually zero to simplify physical interpretation.

Typical is the case of $l_{max} = 1$ for hydrogens, $l_{max} = 3$ for the second row elements in the periodic table, and $l_{max} = 4$ for the rest. The model cannot be easily extended to heavy elements because the valence shell is very thin and high angular orbitals delocalize the charge making it nearly spherical, so high order multipoles are needed to describe correctly the density. Thus, the physical interpretation of multipolar terms becomes difficult. The model reaches its limits. κ' is analogous to κ.

The ratio parameters to fit over data available is too high. Using the multipolar model 27 or 36 parameters have to be fitted for each non hydrogen atom. *High order refinement* is necessary to avoid correlation of thermal factors and multipolar parameters. They recover different types of anisotropy.

The great benefit of this model is that phases are also improved.

The expansion given by the multipolar model for the density is analogous to the expansion by the LCAO formalism, but all terms are atom-centered [1]. The multipolar model has limitations to describe diffuse electronic distributions, $\rho(\mathbf{r}) \approx 0$, such as interatomic regions, due to the nucleus-centered nature mentioned above.

1 Remember that the density resulting from the LCAO formalism expands the wave function in one and two center terms.

2.6.4 Density maps

Deformation maps, $\Delta\rho(r) = \rho(r) - \rho_0(r)$, reveal redistributions of the electron density with respect to spherical non interacting isolated atoms, mapped on a suitable molecular plane. Density arrangements appear at bonds, lone pairs. For a long time density maps were the main source of chemical information [38], but now they have been replaced by QTAIM analysis. Nevertheless, density maps are a good assessment tool to decide the quality of the refinement.

2.7 OTHER EXPERIMENTAL OBSERVABLES

The electrostatic potential is the main byproduct of the density we are interested in. It can be obtained directly from the structure factors or from the total density [2]

$$\phi(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{H}} \frac{F(\mathbf{H})}{\mathbf{H}^2} e^{-2\pi i \mathbf{H}\mathbf{r}} = \int \frac{\rho_t(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' = \sum_{\mathbf{R}} \int_\Omega \frac{\rho_t(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'+\mathbf{R}|}\,d\mathbf{r}' \quad (2\text{–}12)$$

The inconveniences of these expansions are that the reciprocal space series is not defined at $\mathbf{H} = 0$, and the direct space sum is conditionally convergent.

If the unit cell has neutral charge $\int_\Omega \rho_t(\mathbf{r})\,d\mathbf{r} = 0$ and no dipole, converting the sum over direct lattice vectors to a sum over spherical shells [104] transforms it into an absolutely convergent sum. With an Ewald like separation of the integration range in two ranges delimited by a sphere of radius μ^2, the issues of direct and reciprocal space are overcomed to get the best of both. Then, the electrostatic potential is

$$\phi(\mathbf{r}) = \int_\Omega \left(G(\mathbf{r},\mathbf{r}') - \frac{2\pi}{3\Omega}\mathbf{r}'\cdot\mathbf{r}' \right) d\mathbf{r}'. \quad (2\text{–}13)$$

where

$$G(\mathbf{r},\mathbf{r}') = \sum_{\mathbf{R}} \frac{\operatorname{erfc}(\mu|\mathbf{r}-\mathbf{r}'+\mathbf{R}|)}{|\mathbf{r}-\mathbf{r}'+\mathbf{R}|} + \frac{4\pi}{\Omega}\sum_{K\neq 0} \frac{e^{-k^2/4\mu^2}}{k^2} e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}. \quad (2\text{–}14)$$

Standard experiments cannot provide an electrostatic potential reference, which has to be understood as arbitrary. Nevertheless, since we are basically interested in relative differences and topological features do not depend on rigid translations of scalar fields, this is not such an important issue. In any case, the zero potential can be obtained from electron diffraction, but that technique usually destroys soft molecular crystals, because electron beams interact more strongly with matter than X-rays. That makes it amenable to experimental determination only for ionic compounds. So, we usually don't talk about the

[2] knowing that the density has translational symmetry $\rho_t(\mathbf{r}) = \rho_t(\mathbf{r}+\mathbf{G})$.

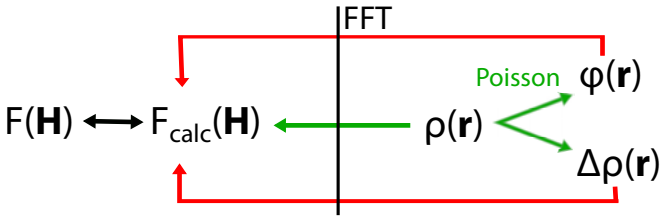


Figure 2: F(H) are the experimental structure factors. The experimental density is the one that best regenerates the experimental structure factors. Both are related by a Fourier transform. The Laplacian is obtained with numerical derivation of the density, and the electrostatic potential by using Poisson's equation.

electrostatic potential of the solid, but instead we take a moiety of the crystal and from the multipolar parameters associated to the atoms in the moiety we obtain its isolated electrostatic potential. The expression is similar to the multipolar expansion and the evaluation is analytic. We are analyzing pseudo-isolated molecules [3], however, the parameters are those of the crystal. The short-range part of the electrostatic potential from the multipolar model applying the scheme of (2–14) is [4]

$$\phi_{sr} = \sum_{A}\sum_{i}\sum_{l=0}^{l_{max}}\sum_{m=-l}^{l} C_{lm}^{Ai} Y_{lm}^{*}(\theta,\varphi) \sum_{\mathbf{R}} \int r_A^2 R_l^{Ai}(r_A) F_l(|\mathbf{r}-\mathbf{R_a}+\mathbf{R}|, R_A, \mu)\, dr_A. \tag{2–15}$$

where

$$F_l(R,r,t) = (2N+1)\int_0^t \frac{e^{(R^2+r^2)/4\tau}}{2\sqrt{\pi\tau^3}} i_l\left(\frac{rR}{2\tau}\right) \tag{2–16}$$

and $i_l(z)$ is the modified spherical Bessel function of first kind.

As well, the electric field, electron density gradient, and Laplacian of the electron density may be obtained through their corresponding Fourier transformation of the structure factors. Here, the completeness of the Fourier series is critical for the density gradient and Laplacian, the latter being

$$\nabla^2\rho(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{H}} H^2 F(\mathbf{H}) e^{-2\pi i \mathbf{H}\mathbf{r}}. \tag{2–17}$$

So, instead it is evaluated by numerical differentiation of the multipolar density.

2.8 ELECTRONIC STRUCTURE OF SOLIDS

One of of the first challenges that quantum mechanics faced was the explanation of metal properties. The simple change of Boltzman statistics for a Fermi-

3 Pseudo-isolated analysis is a common crystallographic procedure to focus on the interactions in specific parts of the crystal.

Dirac distribution of velocities in the classic Drude model has strong consequences on paramagnetism, thermal and electric conductivity, as discussed by Pauli and Sommerfeld. Notwithstanding, the underlying free electron model seemed to Felix Bloch an oversimplified model of metals. Pursued by the impetus of the new emerging theory, Bloch scored a bullseye showing how the eigenstates of a particle under a periodic potential should be.

2.8.1 A toy model of one electron per cell

For a one-dimensional toy model of one particle per cell in a periodic potential $V(\mathbf{r}) = V(\mathbf{r}+\mathbf{R})$ there is no physical justification to think that the electron density will be different from cell to cell.

The fact that the Hamiltonian is invariant to lattice translation operations (commutes with all translation operators that displace the coordinates by a vector of the lattice) and translation operators commute among them constitutes enough evidence to say that all operators can share a simultaneous common eigenbasis set $\{\psi_n\}$

$$H\psi_n = \varepsilon_n\psi_n, \qquad T_{\mathbf{R}}\psi_n = \tau(\mathbf{R})\psi_n. \tag{2–18}$$

Translations have to preserve normalization,

$$|\psi_n(\mathbf{r})|^2 = |\psi_n(\mathbf{r}+\mathbf{R})|^2 = |T_{\mathbf{R}}\psi_n(\mathbf{r})|^2 = |\tau(\mathbf{R})\psi_n(\mathbf{r})|^2 = |\tau(\mathbf{R})|^2|\psi_n(\mathbf{r})|^2, \tag{2–19}$$

thus the wavefunction is allowed to change only in a phase factor. Any combination of translations complies with the identity $\tau(\mathbf{R}+\mathbf{R}') = \tau(\mathbf{R})\tau(\mathbf{R}')$. Irremediably leads to $\tau(\mathbf{R}) = e^{ik\cdot\mathbf{R}}$. The index n indicates the principal quantum number or excitation level of the state.

The continuum of translational invariances in the free electron model is translated to a discrete translational invariance given by the vectors of the lattice. In other words, momentum was preserved everywhere and now only between equivalent points in different cells. Following the tradition of quantum theory the wavefunctions are labeled by the crystal momentum $\mathbf{k}$. That is, a Bloch state is

$$\psi_{nk}(\mathbf{r}) = e^{ik\cdot r}u_{nk}(\mathbf{r}) \tag{2–20}$$

a plane wave modulated by a periodic function $u_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r}+\mathbf{R})$. The periodic functions are an orthonormal basis

$$\frac{1}{\Omega}\int_{\Omega} u^*_{nk}(\mathbf{r})u_{pk}(\mathbf{r})\,d\mathbf{r} = \delta_{np} \tag{2–21}$$

inside the cell. The role of $\mathbf{k}$ in this crystal should not be confused with its counterpart in a free electron gas due to ψ_{nk} not being eigenfunctions of the momentum operator

$$-i\nabla\psi_{nk}(\mathbf{r}) = e^{ik\cdot r}(\mathbf{k}-i\boldsymbol{\nabla})u_{nk}(\mathbf{r}). \tag{2–22}$$

Bloch's theorem maps the problem of expressing the wavefunction in terms of an infinite number of electrons to an infinite number of $\mathbf{k}$-points in the 1BZ. To alleviate the problem the points are discretized in a grid.

Generalization to many-electron wavefunctions is somewhat analogous.

2.8.2 Band structure

Formulated in the reciprocal space, the Schrödinger equation

$$H_{\mathbf{k}} u_{n\mathbf{k}} = \varepsilon_{n\mathbf{k}} u_{n\mathbf{k}} \qquad H_{\mathbf{k}} = (\mathbf{k} - i\boldsymbol{\nabla})^2 + V \tag{2–23}$$

does not change when applying reciprocal lattice translations, because the boundary conditions allow us to state that

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \psi_{n\mathbf{k}+\mathbf{G}}(\mathbf{r}), \tag{2–24}$$

so we can bound the $\mathbf{k}$ points to be in the irreducible Brillouin zone (1BZ). There are infinite $\mathbf{k}$ points in the 1BZ that lead to an infinite number of parametrized Hamiltonians, therefore an infinite of continuous energy values for a band n dispersed in a range of energy values that may overlap with other bands or not.

Being the case a non-interacting system of electrons, the Hamiltonian is a one body operator that can act on $u_{n\mathbf{k}}$ as well as on the first order density matrix. The latter is not so different from the stated expression for a determinantal wavefunction. It is a sum over the occupied band states n with the addition of being integrated in the 1BZ. Bands are filled in increasing order of energies in the ground state and each one can be populated at most with two electrons per cell (spinless model).

$$\gamma_1(\mathbf{r}, \mathbf{r}') = \frac{2\Omega}{8\pi^3} \sum_n \int_{\Omega^*} \theta(\varepsilon_F - \varepsilon_{n\mathbf{k}}) \psi_{n\mathbf{k}}(\mathbf{r}) \psi^*_{n\mathbf{k}}(\mathbf{r}')\, d\mathbf{k} \tag{2–25}$$

where θ is a step function with value 1 when its argument is positive, 0 otherwise. The Fermi energy ε_F is a parameter controlling that the sum is only over occupied states. Its definition arises when a variable dependent $\gamma_1(\varepsilon)$ is integrated in the unit cell

$$N(\varepsilon) = \frac{2\Omega}{8\pi^3} \sum_n \int_{\Omega^*} \theta(\varepsilon - \varepsilon_{n\mathbf{k}})\, d\mathbf{k} \tag{2–26}$$

and forced to be equal to the number of electrons per cell $N(\varepsilon_F) = N$.

The ground state energy of this spinless model is

$$E = \frac{2\Omega}{8\pi^3} \sum_n \int_{\Omega^*} \theta(\varepsilon_F - \varepsilon_{n\mathbf{k}}) \varepsilon_{n\mathbf{k}}\, d\mathbf{k}. \tag{2–27}$$

When the maximum energy (like HOMO) value in the full range of $\mathbf{k}$ of the band $\varepsilon_{N/2}$ is separated from the minimum energy (like LUMO) over $\mathbf{k}$ of the

next band by a gap greater than zero the system is an insulator. An amount of energy greater or equal to the gap separating the bands is needed to excite electrons from valence bands to conducting bands. If the highest occupied band is not full the excitation only requires a very small supply of energy to have conducting electrons.

In a broader sense band properties are more amenable to be discussed in terms of the density of states

$$g(\varepsilon) = dN(\varepsilon)/d\varepsilon = \frac{2\Omega}{8\pi^3} \sum_n \int_{\Omega^*} \delta(\varepsilon - \varepsilon_{n\mathbf{k}})\, d\mathbf{k} = \sum_n g_n(\varepsilon). \tag{2–28}$$

2.8.3 Plane waves and pseudopotentials

Electrons treated in a mean-field approximation have a wavefunction that is an antisymmetrized product of Bloch states $\{|\psi_{n\mathbf{k}}\rangle\}$ with eigenenergies $\varepsilon_{n\mathbf{k}}$. The most obvious choice to expand the periodic part of a Bloch statefunction is to use plane waves. In favor they have that the kinetic operator is diagonal, orthogonality of the basis vectors rules out the clinging basis-set superposition error, fast algorithms exist, and they are already Bloch functions. Much to our regret, the regions of high electronic concentration involve a huge number of basis functions to reach a good approximation of their oscillations, in opposition to the Gaussian basis-sets that are present in most molecular calculations.

Apparently such problems to accurately describe the nuclear region with plane waves do not hinder the success that the free electron model has achieved modelling electrons in simple metals. However, if we insist in treating equally crystals made of atoms with a large number of electrons, where not all electrons become conducting, the description of valence states by everywhere smooth functions (plane waves) fails at the now not so small nuclear regions. Simply shortcutting the wavefunction to contain only valence states and applying the same Hamiltonian has the undesired effect of converging to core states. In consequence, we would need a large number of functions, the first ones would converge to core states and only after all core states have been filled would valence states be represented. The method of Orthogonalized Plane Waves (OPW) [80] prevents this issue orthogonalizing core and valence states as we will see below.

The eigenstates of the mean-field Hamiltonian are conveniently categorized in valence ψ_v and core ψ_c states each with eigenvalues ε_v and ε_c. All-electron valence states are expanded as a linear combination

$$|\psi_v\rangle = |\psi_v^{PP}\rangle - \sum_c \langle\psi_c|\psi_v^{PP}\rangle|\psi_c\rangle \tag{2–29}$$

of a smooth valence function ψ_v^{PP} and core electron functions ψ_c, the latter assumed to be known. The purpose is to have new valence functions that bypass the oscillations inside the atomic core regions. Meanwhile they are intended to be identical to the all-electron valence functions outside. Albeit pseudo-valence

functions are not orthogonal the coefficients of the core expansion ensure a Gram-Schmidt orthogonalization between all-electron valence and core functions.

Examined more carefully, the imposition of orthogonality to avoid valence electrons collapse onto core states is the same as saying that valence electrons do not feel the true nuclear potential in nuclear regions, as it is somehow screened by inner electrons, so that the potential that valence electrons feel is smooth enough to let the valence electrons be described by plane waves. OPW can be recast [150] as the equations for valence states with an effective potential

$$H\psi_v^{PP} - \sum_c \langle\psi_c|\psi_v^{PP}\rangle H\psi_c = \varepsilon_v\psi_v^{PP} - \sum_c \langle\psi_c|\psi_v^{PP}\rangle\varepsilon_v\psi_c \tag{2–30}$$

$$\left[H + \sum_c (\varepsilon_v - \varepsilon_c)|\psi_c\rangle\langle\psi_c|\right]|\phi_v^{PP}\rangle = \varepsilon_v\phi_v^{PP}$$

$$(H + V_R)|\phi_v^{PP}\rangle = \varepsilon_v|\phi_v^{PP}\rangle. \tag{2–31}$$

The new term V_R takes care of the effect of the core electrons on the valence electrons, altering the potential so that $V_{pseudo} = V_R + V$. Since core states are lower in energy than valence states V_R is repulsive mimicking the disturbance produced by core electrons. The pseudopotential can be precomputed but care has to be taken because it is dependent on the basis. Heavy atoms susceptible to relativistic effects can be treated with more sophisticated methods without affecting the efficiency when calculating the valence electronic structure. Thus, in one shot we have avoided the oscillations at the core region, reduced the number of states to simulate, and added an easy scheme to include relativistic effects.

The non-uniqueness of the pseudopotential operator has fomented the appearance of several types. The most important ones: conserve the total charge inside a given radius, these are *norm-conserving pseudopotentials*; try to reduce the number of plane wave coefficients in the pseudo-valence expansion, *ultrasoft pseudopotentials*; or intend to map the complete Hilbert-space, formed by all electronic states to a pseudo-Hilbert space with only pseudo-valence states, that is easier to treat, denominated *projected augmented wave pseudopotentials*.

We return now to the use of plane waves. Statefunctions written as a Bloch function that expand the periodic part $u_{nk}(\mathbf{r})$ with plane-waves through a Fourier series

$$\psi_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} = \left(\sum_{\mathbf{G}} u_{nk}(\mathbf{G})e^{i\mathbf{G}\cdot\mathbf{r}}\right)e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{\mathbf{G}} u_{nk}(\mathbf{G})e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tag{2–32}$$

with coefficients $u_{nk}(\mathbf{G})$. A remarkable feature of plane waves is that unlike for Gaussian basis-set expansions the accuracy is here monotonically improved with the size of the basis. This means we can truncate the expansion to be less than a cutoff parameter

$$\frac{1}{2}|\mathbf{k}+\mathbf{G}|^2 < E_{cut} \tag{2–33}$$

beyond which we have previously checked that no improvement is appreciated. The kinetic operator is diagonal with entries $T_{\mathbf{k}}(\mathbf{G}) = \frac{1}{2}(\mathbf{k}+\mathbf{G})^2$ so E_{cut} is the maximum allowable kinetic energy that a plane-wave can have.

Wavefunctions of insulators or semiconductors vary smoothly from point to point in the reciprocal space so a rapid convergence is achieved with weighted average sums over special points designed after the lattice point group of the crystal. The most widely used grid, a Monkhorst-Pack grid [140], for integrating properties is derived from the special point integration. Metals can pose a challenge because the occupied bands can vary a lot from one point to other, being the Fermi surface a complicated manifold that destroys the beauty of the averaging process. To maintain a good average, many times the Fermi surface itself is smoothened, replacing the Heaviside function with some smearing function like a Gaussian spreading, Methfessel-Paxton [128] or the Fermi-Dirac distribution. For metals, the tetrahedron method [23] is often advised.

The most direct approach to calculate the electronic structure of a solid is to employ DFT, whereas the extrapolation of HF for periodic systems is not as straightforward. Within DFT an initial density, for example from a superposition of atomic densities, is provided, and upon it the selected exchange-correlation functional is applied. The external potential is divided in ionic contributions with a form that depends on the pseudopotential chosen. The local part is better calculated in direct space since it is more or less diagonal. On the other hand, the computation of the long-range part is more efficient in direct space when the unit cell is big, otherwise is done in reciprocal space. After all evaluations in direct space have been performed the terms are transformed to reciprocal space through Fourier transforms. The density and wavevectors

$$u_{nk}(\mathbf{G}) = \frac{1}{\Omega}\int_{\Omega} e^{-i\mathbf{G}\cdot\mathbf{r}} u_{nk}(\mathbf{r})\,d\mathbf{r} \tag{2–34}$$

$$n(\mathbf{G}) = \frac{1}{\Omega}\int_{\Omega} e^{-i\mathbf{G}\cdot\mathbf{r}} n(\mathbf{r})\,d\mathbf{r} \tag{2–35}$$

$$\tag{2–36}$$

are restricted to a maximum cutoff $\frac{1}{2}|\mathbf{G}|^2 \leqslant 4E_{cut}$ to be in line with the kinetic energy cutoff. Then, the kinetic energy and the remaining parts of the potential can be calculated so we are ready to evaluate the Schrödinger equation

$$\sum_{\mathbf{G}'} H_{\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}'} c_{n,\mathbf{k}+\mathbf{G}'} = \varepsilon_n c_{n,\mathbf{k}+\mathbf{G}'} \tag{2–37}$$

with matrix elements

$$H_{\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}'} = \frac{1}{2}|\mathbf{k}+\mathbf{G}|^2\delta_{\mathbf{G},\mathbf{G}'} + V_{ion}(\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}') + V_H(\mathbf{G}-\mathbf{G}') + V_{XC}(\mathbf{G}-\mathbf{G}'). \tag{2–38}$$

Returning back to direct space is easy with the Fourier transform. The new density is

$$\rho(\mathbf{r}) = \frac{2\Omega}{8\pi^3} \sum_n \int_{\Omega^*} \theta(\varepsilon_F - \varepsilon_{n\mathbf{k}}) |\psi_{n\mathbf{k}}(\mathbf{r})|^2 \, d\mathbf{k}. \tag{2–39}$$

We repeat from the beginning until convergence in the energy and wavefunctions is reached.

Plane-waves plus pseudopotential calculations can be very incredibly efficient and accurate predicting the energies but reconstruction of the density can be problematic. Indeed, pseudopotentials are not the only choice. All-electron methods use a mixed basis, a Gaussian basis set inside the core region defined by a radius and plane-waves at interstitial regions. So, instead of freezing electrons this time the full electronic structure is simulated with the consequent computational cost.

2.8.4 Projected Augmented Waves (PAW)

Those of you interested in reconstructing the density functions from the pseudo-valence orbitals and the core pseudo-cation orbitals without resorting to all-electron methods can adopt the projection technique of plane augmented waves (PAW) [22]. PAW unifies the pseudopotential and the all-electron method of linearized augmented plane waves based on the assumption that the Hilbert space of all the valence states (Ψ) orthogonal to the core states and the Hilbert space of the pseudo-valence states ($\tilde{\Psi}$) are related by a linear transformation

$$\Psi = \mathcal{T}\tilde{\Psi} \tag{2–40}$$

that relates operators in both spaces through

$$\langle \Psi | A | \Psi \rangle = \langle \mathcal{T}\tilde{\Psi} | A | \mathcal{T}\tilde{\Psi} \rangle = \langle \tilde{\Psi} | \mathcal{T}^* A \mathcal{T} | \tilde{\Psi} \rangle = \langle \tilde{\Psi} | \tilde{A} | \tilde{\Psi} \rangle. \tag{2–41}$$

Also, the KS energy accepts a variational formulation in the pseudo-Hilbert space.

Of the many possible linear transformations we can guess that the appropriate ones will mostly affect nuclear regions and be close to identity in interstitial regions. Thus the ansatz

$$\mathcal{T} = 1 + \sum_{R_A} \mathcal{T}_{R_A} \tag{2–42}$$

where R_A are the nuclear regions. There, the true valence wavefunction finds a expansion in terms of true partial waves

$$\Psi = \sum_\nu c_\nu \psi_\nu \tag{2–43}$$

and since pseudo-plane waves are related to the former by $\psi_\nu = \mathcal{T}\psi_\nu^{PP}$,

$$\Psi = \sum_\nu c_\nu \mathcal{T}\psi_\nu^{PP} = \mathcal{T} \sum_\nu c_\nu \psi_\nu^{PP} = \mathcal{T}\tilde{\Psi}, \tag{2–44}$$

both expansions share the same coefficients c_ν. Therefore the difference in core regions has to be

$$\sum_\nu c_\nu(\psi_\nu - \psi_\nu^{PP}). \tag{2–45}$$

The corresponding operator

$$\mathcal{T} = 1 + \sum_\nu \left(\psi_\nu - \psi_\nu^{PP}\right) \langle p_\nu| \tag{2–46}$$

needs a set of projectors p_ν that when applied to $\tilde{\Psi}$

$$\Psi = \mathcal{T}\tilde{\Psi} = \tilde{\Psi} + \sum_\nu \left(\psi_\nu - \psi_\nu^{PP}\right) \langle p_\nu|\tilde{\Psi}\rangle \tag{2–47}$$

project with the coefficients of the true wavefunction, $c_\nu = \langle p_\nu|\tilde{\Psi}\rangle$. The projectors, dual to the pseudo-plane waves ψ_ν^{PP}, $\langle p_i|\psi_j^{PP}\rangle = \delta_{ij}$, have to be generated by some orthogonalization procedure.

Expectation values of operators in terms of the pseudo-waves

$$\langle\tilde{\Psi}|\tilde{O}|\tilde{\Psi}\rangle = \langle\tilde{\Psi}|O|\tilde{\Psi}\rangle + \sum_{i,j}\langle\tilde{\Psi}|p_i\rangle \left(\langle\psi_i|O|\psi_j\rangle - \langle\psi_i^{PP}|O|\psi_j^{PP}\rangle\right) \langle p_j|\tilde{\Psi}\rangle. \tag{2–48}$$

The real space operator $O = |\mathbf{r}\rangle\langle\mathbf{r}|$ can be used to obtain the electron density

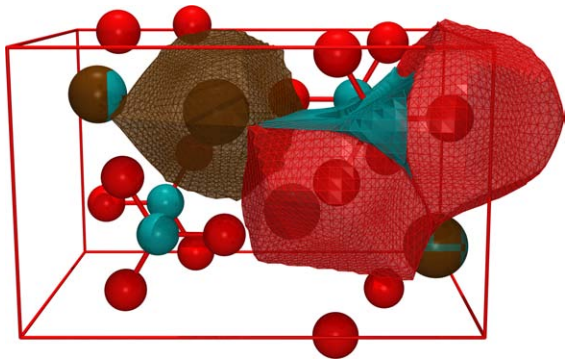
$$\rho(\mathbf{r}) = 2\sum_n^N \left[|\psi_\nu^{PP}(\mathbf{r})|^2 + \sum_{i,j}\langle\psi_\nu^{PP}|p_i\rangle D_{ij}(\mathbf{r})\langle p_j|\psi_\nu^{PP}\rangle\right] \tag{2–49}$$

with functions

$$D_{i,j}(\mathbf{r}) = \psi_i(\mathbf{r})\psi_j(\mathbf{r}) - \psi_{\nu,i}(\mathbf{r})\psi_{\nu,j}(\mathbf{r}). \tag{2–50}$$

3 THEORY OF CHEMICAL BONDING

CONTENTS



Atomic domains in the Aragonite crystal.

3.1 CHEMICAL INNERMOST CONUNDRUMS FOR QUANTUM MECHANICS

Since its origin chemistry has been an empiric science. Chemists had to deal with entities that were out of their reach, tiny particles beyond observation, so they elaborated mechanistic descriptions that agreed with phenomenological observations. The concepts distilled out of elaborated works by Lavoisier, Dalton, Pauli, ... have been always mere devices to embody the corpus of knowledge in a consistent way but today it is impossible to conceive chemistry without bonds, electronegativity, lone pairs, acidity or functional groups. The rise of quantum theory in the beginning of the twentieth century shocked many chemists with an outrageous interpretation of matter. None of the previous concepts found a natural placement in the theory.

The picture of electrons rushing between atoms (chemical arrows) with a defined trajectory is not possible. Distinction of core and valence electrons vanishes, all electrons are indistinguishable now and contribute equally to bonding. Moreover, the reasoning of wave mechanics is not as straightforward as it was with classical mechanics. For example, can we map our physical interpretation of the covalent bond, sharing of electrons between atoms, with an energy lowering? If we try to be consistent with the virial theorem we could guess that there is a lowering of potential energy accompanied by an increase of kinetic energy but if we attend to the spread of electrons we would be driven to the opposite conclusion, therefore violating the virial theorem. Definitively, both are extrema cases of a more realistic compromise, found by variational minimization.

The construction of the wavefunction as a product of orbitals reintroduced the distinction of electrons and a link with the old electronic theory. The bridge was broadened specially after Pauli connected the theory of Heitler-London with Lewis formulation of covalent bonding. A long record of successes (Woodward-Hoffman rules, Hückel's rule, HOMO-LUMO reactivity, ...) explains the presence of molecular orbitals as common objects in organic chemistry. Despite of it, their choice is arbitrary. Equivalent classes of orbitals can represent electrons with diverse spatial localizations, changing completely the interpretation. When a quantitative analysis is needed problems appear because they are not correlated images of one electron distributions.

The message that is left is that chemists need real space representations of the electronic distribution that are compatible with the dictates of quantum mechanics. This road has been started in the 60's and still continues being developed. In this chapter we will tell how quantum mechanics has included those concepts in its glossary and how it requires upon us a proper definition of atoms in molecules.

3.2 ELECTRON POPULATION AND BONDING INDICES ON A PROBABILISTIC FOOTING

In the previous chapters we have shown the fundamentals of quantum methods that solve the non–relativistic Schrödinger equation to reach the best approximation to the wavefunction of the system. Henceforth, we can obtain any measurable quantity as an expectation value of a proper operator. Our objective is to give explanations for chemical concepts based on the electronic distribution and ultimately distinguish and predict bonding types.

When analyzing the wavefunction, one often focuses on the bond distances, total energy or vibrational frequencies to get insight. However, very limited information is contained in those values. It would be a waste of time to compile the full wavefunction for later dismissing most of the information contained in it knowing that all is coded in it. A most important observable for chemists is the electron density because electron are the main characters in atomic interactions. The density of a monodeterminantal wavefunction whose molecular orbitals are expanded as a LCAO is expressed in terms of the basis orbitals as

$$\rho(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu}\chi^*_\mu(\mathbf{r})\chi_\nu(\mathbf{r}) \tag{3–1}$$

Single determinant wavefunctions

The spin-less density matrix P obtained in Hartree–Fock theory is not idempotent, because the basis where it is represented is not orthonormal (primitive or AOs if minimal basis) as it is in the basis of best molecular orbitals. The choice of a MOs basis to represent the density matrix would destroy our purpose because they are delocalized over the whole space and the association to atoms is blur. The matrix PS would be a better choice, $N = \sum_{\mu\nu} P_{\mu\nu}S_{\nu\mu}$. From this density matrix representation we can divide the electron population in atomic contributions. Counting separately the atomic orbitals associated to each atom, the average is

$$N_A = \sum_{\nu\in A} (PS)_{\nu\nu}. \tag{3–2}$$

The populations determined this way are called Mulliken's atomic populations. A drawback is that they might not be positive. Pairs, triads, quartets and so on are as well accessible taking advantage of the idempotency $(PS)^2 = PS$

$$N_A = \sum_{\nu\in A} (PS)_{\nu\mu}(PS)_{\mu\nu} + \sum_{\substack{\nu\in A \\ \mu\notin A}} (PS)_{\nu\mu}(PS)_{\mu\nu}. \tag{3–3}$$

The first term is orbital net population and the second is overlap population, connected with the non-orthogonality of the basis. Indices of populations that distribute among more than two orbitals are not equal because the matrix PS lacks symmetry, Löwdin's orthogonalization is a way to symmetrize Mulliken population matrix

$$N_A = \sum_{\nu\in A} (S^{1/2}PS^{1/2})_{\nu\nu}. \tag{3–4}$$

The use of the pseudo density matrix PS is useful because the orbitals used to represent it can be easily mapped to atomic components. In this formulation the choice of the orbitals is not unique as any unitary transformation (a rotation with respect to the laboratory frame) of the orbitals leaves invariant the wavefunction. The matrix P is therefore modified in the process. The significance of individual elements of the P matrix (configuration space objects) lack physical meaning. Luckily Mulliken analysis remains invariant but Löwdin populations are only invariant under certain rotations. We should seek another mathematical representation of n-RDM objects where discretization of the system in atomic elements meets the requirement of invariance under unitary transformations.

3.2.1 Electron Population Distribution Functions

The difficulties that arise when one seeks a connection between chemistry and quantum mechanics are due to the fact that state functions strictly belong to a $\mathbb{R}^{3N}$ configuration space, in contrast to standard chemical concepts, usually envisioned in the $\mathbb{R}^3$ physical space. In terms of the theory of probability, each configuration is an event worth of consideration, but when we are interested in a particular chemical event only states located in a fragment of the physical space, where chemical action takes place, are relevant. On taking this path we transform the problem of partitioning the orbital space to dividing the physical space into the best set of finite volumes.

The number of electrons in an atom or bond, designed as $\Omega \subset \mathbb{R}^3$, with characteristic function ω_Ω, is an observable that can be obtained from a particle number operator [165]

$$\hat{N}_\Omega = \sum_{i=1}^{N} \omega_\Omega(\mathbf{r}_i) \qquad \omega_\Omega(\mathbf{r}_i) := \begin{cases} 1 & \text{if } \mathbf{r}_i \in \Omega \\ 0 & \text{if } \mathbf{r}_i \in \Omega^c \end{cases} \tag{3–5}$$

This operator does not commute with the Hamiltonian, so the local population of Ω fluctuates around an average particle population

$$N_\Omega = \langle \Psi | \hat{N}_\Omega | \Psi \rangle = \sum_{i}^{N} \int_\Omega d\mathbf{x}_i \int_{\mathbb{R}^3} d\mathbf{x}_{j\neq i} \, |\Psi|^2 = \int_\Omega \rho(\mathbf{r}) \, d\mathbf{r}, \tag{3–6}$$

with variance

$$\sigma^2_\Omega = \langle \Psi | \hat{N}^2_\Omega | \Psi \rangle - \langle \Psi | \hat{N}_\Omega | \Psi \rangle^2 = \int_\Omega \rho(\mathbf{r}) \, d\mathbf{r} + \int_\Omega \rho_2(\mathbf{r}_1, \mathbf{r}_2) \, d\mathbf{r}_1 \, d\mathbf{r}_2 - N^2_\Omega \tag{3–7}$$

determined using the expression of one- and two-body operators in terms of density matrices for

$$\hat{N}^2_\Omega = \left[\sum_{i=1}^{N} \omega_\Omega(\mathbf{r}_i) \right]^2 = \sum_{i=1}^{N} \omega^2_\Omega(\mathbf{r}_i) + \sum_{j\neq i}^{N} \omega_\Omega(\mathbf{r}_i) \omega_\Omega(\mathbf{r}_j). \tag{3–8}$$

A fundamental unanswered question is: what "do electrons do"? From their point of view there are two primitive options: either entering Ω or stay outside. Intermediate actions include electrons being in Ω a fraction of the time π_Ω and $1-\pi_\Omega$ of the time outside. Their attitude is ultimately driven by the election that is more beneficial in energetic terms for the whole system. A correlated equilibrium. Non-commutativity says that electrons cannot have a pure strategy, determined by a well defined condensation, e.g. exactly n electrons inside and $N-n$ outside. Instead, a mixed strategy where at least some (indistinguishable) electrons are indecisive to take the action of localizing in Ω or Ω^c is the experimentally observed choice.

Quantification of the probability for each pure strategy is accomplished with

$$p(n \in \Omega, N-n \in \Omega^c) \equiv p_n(\Omega) = \binom{N}{n} \int_\Omega \mathrm{d}\boldsymbol{x}_{[1,n]} \int_{\Omega^c} \mathrm{d}\boldsymbol{x}_{(n,N]} \, |\Psi|^2. \tag{3–9}$$

A combinatoric number of indiscernible electron condensations add up with the same weight. Satisfies normalization $\sum_{n=0}^{N} p_n = 1$.

The probability is also expressed as a sum of coarse-grained integrations of reduced density functions. The smallest reduced density function that can describe a n electron condensation in Ω is obviously ρ_n. ρ_n is the result of integrating $n+1$ to N electrons in all space. Notice that integration domains are different. There is no constraint, as in (3–9), to the complementary space of Ω. Therefore, integration of ρ_n in Ω gives an overestimation if $N > 1$. This is easily seen when $\Omega = \mathbb{R}^3$: $\int_\Omega \gamma_1 = \operatorname{Tr} \gamma_1 = N$. Next, subtracting the integration of $\int_\Omega \int_\Omega \gamma_2$ creates a lower bound because a bunch of triads or more electrons live in Ω, and we have subtracted them off more than once. Successive repetition of this procedure until ρ_N eventually gives the right answer. The correct formula applying the inclusion-exclusion principle is

$$p_n(\Omega) = \frac{1}{n!} \sum_{i=0}^{N-n} \frac{(-1)^i}{i!} \int_\Omega \rho_{n+i}(\boldsymbol{x}_{[1,n+i]}) \, \mathrm{d}\boldsymbol{x}_{[1,n+i]}. \tag{3–10}$$

More general probabilities are also the result of linear combinations of integration of reduced density functions.

Further insight on the probabilities can be gained by building the distribution from scratch assuming independence. Let the independent random variable Π_i, related to an electron i falling ($\Pi_i = 1$) or not ($\Pi_i = 0$) inside Ω, follow a Bernoulli distribution with parameter π_i. The number of electrons condensed will be the sum of N independent events $n = \sum_i^N \Pi_i$. To decipher the distribution we use the probability generating function to exploit the assumed independence:

$$G(t) = \langle t^n \rangle = \prod_{i=1}^{N} \langle t^{\Pi_i} \rangle = \prod_{i=1}^{N} \sum_{k=0}^{1} t^k \pi_i^k (1-\pi_i)^{1-k} = \prod_{i=1}^{N} (t\pi_i + 1 - \pi_i). \tag{3–11}$$

When the electrons are also identically distributed ($\pi_i \to \pi$)

$$(t\pi + 1 - \pi)^N = \sum_{\nu=0}^{N} t^n \binom{N}{n} \pi^n (1-\pi)^{N-n} = \sum_{n=0}^{N} t^n p_n \tag{3–12}$$

n electron condensations have a binomial distribution. On the other hand, from this result we can infer

$$\begin{aligned} G(t) &= \prod_{i=1}^{N} (t\pi_i + 1 - \pi_i) \\ &= \sum_{n=0}^{N} t^n \sum_{\substack{I_n \subseteq [\![1,N]\!] \\ \mathrm{card}(I_n) = n}} \left(\prod_{i \in I_n} \pi_i \right) \left(\prod_{i \notin I_n} (1-\pi_i) \right) = \sum_{n=0}^{N} t^n p_n \end{aligned} \tag{3–13}$$

a Poisson-Binomial distribution. In the case we knew the parameters, efficient algorithms to compute the distribution would be available. That information will be given by a careful analysis of the wavefunction.

It has been only quite recently that computation of probabilities has become feasible even for the crudest wavefunctions. Starting from a monodeterminantal wavefunction with real spinorbitals, the squared wavefunction is

$$\begin{aligned} |\Psi|^2 &= \left(\det\{\psi_i(\mathbf{x}_j)\}_{i,j=1}^{N} \right)^2 \Big/ N! \\ &= \det \left\{ \sum_{k=1}^{N} \psi_i(\mathbf{x}_k) \psi_j(\mathbf{x}_k) \right\}_{i,j=1}^{N} \Big/ N! \\ &= \sum_{\mathbf{k} \in \{1,\dots,N\}^N} \det \left\{ \psi_i(\mathbf{x}_{k_j}) \psi_j(\mathbf{x}_{k_j}) \right\}_{i,j=1}^{N} \Big/ N! \\ &= \sum_{\mathbf{k} \in \mathcal{S}_N} \det \left\{ \psi_i(\mathbf{x}_{k_j}) \psi_j(\mathbf{x}_{k_j}) \right\}_{i,j=1}^{N} \Big/ N! \,. \end{aligned} \tag{3–14}$$

First, the fundamental property of determinants ($\det(A)\det(B) = \det(AB)$) is used to get a single determinant with elements that are sums of orbital pair products. Next, using the multilinearity of the determinant, columns $\sum_k A_k$, $\sum_k B_k$, etc. are simplified with elementary operations $|\sum_k A_k, \sum_k B_k, \cdots| = \sum_{k,l} |A_l, B_k, \cdots|$. The last sum of (3–14) runs over all permutations in the symmetric group $\mathcal{S}_N$ instead of a N power set $\{1,\dots,N\}^N$ as in the previous equality because when $k_i = k_j (i \neq j)$ the columns are not linearly independent, therefore the determinant is zero.

Further simplification makes use of local overlap matrices, with elements

$$\begin{aligned} S_{ij}(\Omega) &= \int_{\Omega} \psi_i(\mathbf{x})\psi_j(\mathbf{x})\,d\mathbf{x}, \\ S_{ij}(\Omega^c) &= \int_{\Omega^c} \psi_i(\mathbf{x})\psi_j(\mathbf{x})\,d\mathbf{x}. \end{aligned} \tag{3–15}$$

It is understood that if the orbitals are orthonormal the overlap matrix extended to all space is the identity matrix $S(\mathbb{R}^3) = \mathbb{1}_N$. After replacing $|\Psi|^2$ from equation (3–14) inside equation (3–9) we arrive to a formula

$$p_n = \binom{N}{n} \frac{1}{N!} \sum_{\tau \in S_N} \det \left\{ \widehat{S_{i\tau_j}} \right\}_{i,j=1}^{N} \tag{3–16}$$

simplified with

$$\widehat{S_{i\tau_j}} = \begin{cases} S_{i\tau_j}(\Omega) & \text{if } \tau_j \in [1, n] \\ S_{i\tau_j}(\Omega^c) & \text{if } \tau_j \in (n, N]. \end{cases} \tag{3–17}$$

p_n is the sum over all determinants where n columns of $S(\Omega^c)$ have been replaced with columns of $S(\Omega)$. The overlap matrix $S(\Omega)$, apart from commuting with $S(\Omega^c) = \mathbb{1}_N - S(\Omega)$, admits a similarity transformation $D = R^t S R$, employing a unitary matrix R such that $R^t R = \mathbb{1}_N$, into a diagonal matrix D with elements π_i that are the parameters we needed. That relates the Poisson-Binomial probabilities (3–18) with the probabilities of condensing electrons of a determinantal function in bipartite spaces. Recognition of this link allows for taking advantage of the mathematical literature to pick an efficient algorithm. Ending the development of the generating function we arrive to a identity

$$\begin{aligned} G(t) &= \prod_{i=1}^{N} (t\pi_i + 1 - \pi_i) \\ &= \det |(tD + \mathbb{1}_N - D)| \\ &= \det |R^t R(tD + \mathbb{1}_N - D)| \\ &= \det |R^t (tD + \mathbb{1}_N - D) R| \\ &= \det |tR^t D R + R^t R - R^t D R| \\ &= \det |tS + \mathbb{1}_N - S| = \sum_{n=0}^{N} t^n p_n \qquad \forall t \in \mathbb{R}^3 \end{aligned} \tag{3–18}$$

first put in context by Cancès et al. [30], that serves to implement the computation of the probabilities as a linear system of equations in t. Importantly, one can calculate the probabilities with a recurrence formula that scales quadratically.

Once the underlying distribution is known selected domain properties can be easily evaluated in terms of the probability measure p_n. As a result the average population is

$$N_\Omega = \langle n \rangle = \sum_{n=0}^{N} n p_n. \tag{3–19}$$

Remark that an average value of $N_\Omega = k \in \mathbb{N}^*_{<N}$ does not mean, in general, that there are exactly k electrons localized in Ω. Any distribution symmetric around the leading event has the same mean but differing multi body fluctuations. In turn, the dispersion of electron pairs is contained in the variance, recovered this way:

$$\begin{aligned}\sigma^2_\Omega &= \sum_{n=0}^{N} (n - N_\Omega)^2 p_n = \sum_{n=0}^{N} (n^2 - N^2_\Omega) p_n = \sum_{n=0}^{N} (n^2 - n + n - N^2_\Omega) p_n \\ &= \sum_{n=0}^{N} \left[n(n-1)p_n + n p_n - N^2_\Omega p_n \right] = \langle \rho_2 \rangle_{\Omega,\Omega} + N_\Omega - N^2_\Omega = N_\Omega - \lambda_\Omega.\end{aligned} \tag{3–20}$$

Firstly, check out that it is the same as equation (3–7). When the variance is zero there is a pure pair population $\langle \rho_2 \rangle_{\Omega,\Omega} = N_\Omega(N_\Omega - 1)$. The term N_Ω counterbalances the self-interaction present in $\langle \rho_2 \rangle_{\Omega,\Omega}$. From another point of view the process we describe is quantified by counting particles in a subsample of the space. If particles are randomly distributed in the region, the pair population will oscillate the same extent as the average number of electrons ascribed to the volume. Any reduction from this value is a measure of both pair correlation and localization. Certainly, the deviation originated by the correlation of electron pairs contained in Ω, or localization index,

$$\lambda_\Omega = \int_\Omega \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 = \langle \rho_2 \rangle_{\Omega,\Omega} - N^2_\Omega \tag{3–21}$$

that appears after replacing $\rho_2 = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) - \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$, is such quantity. In the limit case of zero variance attains a value of N_Ω [8] meaning that all the pairs of electrons found inside are localized.

Pursuing the idea that covalent bonding is due to electron sharing in real space we can inspect the variance of two domains $\Theta = \Omega \cup \Lambda$ with average populations N_Ω and N_Λ now that we have a deeper insight of the variance from a probabilistic perspective. With n population for Ω and ν for Λ and total population in Θ of $n + \nu$ the joint variance is

$$\begin{aligned}\sigma^2_\Theta &= \mathbb{E}\left[\left([n + \nu] - [N_\Omega + N_\Lambda] \right)^2 \right] = \mathbb{E}\left[\left([n - N_\Omega] - [\nu - N_\Lambda] \right)^2 \right] \\ &= \mathbb{E}\left[\left(n - N_\Omega \right)^2 + \left(\nu - N_\Lambda \right)^2 - 2(n - N_\Omega)(\nu - N_\Lambda) \right] \\ &= \sigma^2_\Omega + \sigma^2_\Lambda - 2\mathrm{cov}(n, \nu)\end{aligned} \tag{3–22}$$

$$\lambda_{\Omega,\Lambda} = -2\int_{\Omega} d\mathbf{r}_1 \int_{\Lambda} d\mathbf{r}_2\, \rho_2(\mathbf{r}_1,\mathbf{r}_2) - N_\Omega N_\Lambda = -2\mathrm{cov}(n,\nu) \tag{3–23}$$

pays attention to the two way fluctuation of charge through the boundary that separates both domains. In other words, it is the charge of Ω delocalized in Λ and vice versa. A symmetric measure of electronic excess for completely general correlated state functions

$$\delta^{\Omega\Lambda} = +\lambda(\Omega,\Lambda) + \lambda(\Lambda,\Omega) = 2\int_{\Omega} d\mathbf{r}_1 \int_{\Lambda} d\mathbf{r}_2\, \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2) \tag{3–24}$$

is known as the delocalization index. During the last decade an increasing number of works have widened the relation that exists between the delocalization index and covalent bonding. On a separate issue, part of our research in chapter 6 has been focused on finding the relation between the DI and electrical conductivity.

Electrons with different spin can be differentiated with a label as alpha or beta electrons: $N = N^\alpha + N^\beta$. Structures consisting in n^α_Ω and $n^\alpha_{\Omega^c}$ α electrons in the domains Ω and Ω^c, $N^\alpha = n^\alpha_\Omega + n^\alpha_{\Omega^c}$; and equally for beta electrons: n^β_Ω and $n^\beta_{\Omega^c}$ β electrons in domains Ω and Ω^c, $N^\beta = n^\beta_\Omega + n^\beta_{\Omega^c}$, are described by a $S \equiv \{n^\alpha_\Omega, n^\alpha_{\Omega^c}, n^\beta_\Omega, n^\beta_{\Omega^c}\}$ nomenclature. Note that this distinction doubles the number of basins to four, two of them for alpha electrons and two for beta electrons.

The average population of alpha electrons in Ω is

$$\begin{aligned} n^\alpha_\Omega &= \sum_{n=0}^{N^\alpha} n p(n^\alpha_\Omega = n) \\ &= \sum_{n=0}^{N^\alpha} n \sum_{x=0}^{N^\beta} p(n^\alpha_\Omega = n, n^\alpha_{\Omega^c} = N^\alpha - n, n^\beta_\Omega = x, n^\beta_{\Omega^c} = N^\alpha - x) \end{aligned} \tag{3–25}$$

runs through all possible n^β_Ω, $n^\alpha_{\Omega^c}$, $n^\beta_{\Omega^c}$.

Hartree–Fock wavefunctions include no correlation among alpha and beta electrons [112], so a factorization in α and β probabilities is due

$$p(n^\alpha_1, n^\alpha_2, n^\beta_1, n^\beta_2) = p^\alpha(n^\alpha_1, n^\alpha_2) \times p^\beta(n^\beta_1, n^\beta_2). \tag{3–26}$$

It is evident that with the only determination of one of them, e.g. $p^\alpha(n^\alpha_1, n^\alpha_2)$, we have complete knowledge of the entire probability.

In principle, all the properties discussed up to now can be extended to an arbitrary decomposition of space into m mutually exclusive regions $\mathbb{R}^3 = \cup_{i=1}^{m} \Omega_i$, $\Omega_i \cap \Omega_j = \emptyset$ $(i \neq j)$. We accordingly fragment the total electron distribution in spatial domains to see how much electronic charge is condensed in them by exploring the probabilities of all possible arrangements of electrons in those domains. If we ascribe the distribution to a finite set of volumes, a

configuration of n_1 electrons in domain Ω_1, n_2 in domain Ω_2, ..., conforming a total of N electrons in m domains, has a probability given by a multiset choice

$$\begin{aligned} p(n_1, n_2, \cdots) &= \sum_{i_1,\ldots,i_N=1}^{N} \omega_{\Omega_{i_1}}(\mathbf{r}_1)\cdots\omega_{\Omega_{i_N}}(\mathbf{r}_N) \sum_{\sigma_1,\ldots,\sigma_N} \int |\Psi|^2 \, d\mathbf{r}_{[1,N]} \\ &= \frac{N!}{n_1!n_2!\cdots n_m!} \int_{\Omega_1} d\boldsymbol{x}_{i\leqslant n_1} \int_{\Omega_2} d\boldsymbol{x}_{(n_1,n_1+n_2]} \cdots \int_{\Omega_m} d\boldsymbol{x}_{(N-n_m,N]} |\Psi|^2. \end{aligned} \tag{3–27}$$

Inter-domain population probabilities are independent for the integrations are multiplied in any order. Correlation is implicit. When all electrons inside basin Ω_i are identically distributed a multinomial distribution

$$p(n_1, n_2, \cdots) = \frac{N!}{n_1!n_2!\cdots n_m!} \prod_{i=1}^{m} \pi_{\Omega_i}^{n_i} \tag{3–28}$$

with unconditional one particle probabilities

$$\pi_{\Omega_i} = \frac{1}{N} \int_{\Omega_i} \rho(\mathbf{r}) \, d\mathbf{r}. \tag{3–29}$$

During this section it has been dropped the idea that the race to achieve the best electronic arrangement can be seen as a game played by electrons, the latter developing a series of strategies to maximize the total payload, the energy. In fact, it is interesting to realize the new perspective that is given of traditional concepts in chemistry without violating the rules of quantum mechanics. The commitment displayed by many chemists to cling to localized electrons can be replaced by a more realistic concept of *strategy*, that is between the otherwise funny and totally delocalized picture defended by those that strictly stick to quantum mechanics and the localized picture. Well then, there might exist a series of predominant strategies, applicable to a single domain or a union of domains — it always depends on the choice of the domain —, that gather the ideas we all are used to in chemistry. One can only wonder about how such common chemical concepts could be encapsulated. In the end, EDFs are destined to enhance our understanding of the optimal strategy followed by electrons to distribute over space.

3.2.2 Domain Averaged Fermi Holes

An analogous scheme to Mulliken's condensation for multi–center indices of a single determinant wavefunction is also possible in real space with reduced density matrices exploiting the Fock-Dirac expansion, here

$$N_\Omega = \int_\Omega \rho(\mathbf{r}) \, d\mathbf{r} = \int_\Omega \int_\Omega \gamma_1(\boldsymbol{x}_1, \boldsymbol{x}_2)\gamma_1(\boldsymbol{x}_2; \boldsymbol{x}_1) \, d\boldsymbol{x}_1 \, d\boldsymbol{x}_2 \tag{3–30}$$

$$+ \int_\Omega \int_{\Omega^c} \gamma_1(\boldsymbol{x}_1, \boldsymbol{x}_2)\gamma_1(\boldsymbol{x}_2; \boldsymbol{x}_1) \, d\boldsymbol{x}_1 \, d\boldsymbol{x}_2. \tag{3–31}$$

For more general wavefunctions the ρ_{XC} includes in addition to the cyclic kernel product of non-diagonal 1RDMs (exchange factor) a correlation factor not obtainable from lower order reduced density matrices, precisely both are contained in the cumulant density for an arbitrary wavefunction $\lambda_2(\mathbf{x}_1,\mathbf{x}_2) \equiv \rho_{xc}(\mathbf{x}_1,\mathbf{x}_2) = \gamma_1(\mathbf{x}_1;\mathbf{x}_2)\gamma_1(\mathbf{x}_2;\mathbf{x}_1)$. In our statistical definition we will regard the exchange contribution as part of the correlation. We have seen in section 1.3.1 recurrence relations that hold for a general wavefunction. Opposite to reduced density matrices the cumulants are extensive and allow a consistent truncation scheme.

Owing to the tight link between cumulants and particle fluctuations of the electron distribution, the first and second components on the right hand side of equation (3–31) are a measure of the variance and covariance of electronic population in Ω and $\Omega \leftrightarrow \Omega^c$. Thus, the interpretation of net and overlap populations in this picture is as fluctuations of the electronic population among spatial domains.

Let us consider the integration of the exchange-correlation density before going with more general cumulant densities. The inner integration can be substituted with the averaged exchange-correlation hole over one domain, $g^{\Omega}(\mathbf{x})$,

$$N_\Omega = \int_\Omega g^{\Omega}(\mathbf{x})\,d\mathbf{x} + \int_\Omega g^{\Omega^c}(\mathbf{x})\,d\mathbf{x} \tag{3–32}$$

that results from a charge weighted exchange-correlation hole, $h_{XC}(\mathbf{x}_1,\mathbf{x}_2)$,

$$g^{\Omega}(\mathbf{x}) = N_\Omega \int_\Omega h_{xc}(\mathbf{x},\mathbf{x}_2)\,d\mathbf{x}_2 = N_\Omega \left(\rho(\mathbf{r}) - \frac{\int_\Omega \gamma_2(\mathbf{x},\mathbf{x}_2)d\mathbf{x}_2}{\int_\Omega \rho(\mathbf{x}_2)d\mathbf{x}_2}\right). \tag{3–33}$$

This quantity, known as Domain Averaged Fermi Hole (DAFH), is an exclusion density that captures the spatial localization of one electron in Ω. It is a very important object for getting insights about bonding in molecules. An exhaustive partition of space permits us recover the density

$$\rho(\mathbf{r}) = \sum_{a=1}^{m} \int_{\Omega_a} g^{\Omega_a}(\mathbf{r})d\mathbf{x}. \tag{3–34}$$

It has been used mainly for single Slater determinants because of the unavailability of $\rho_{XC}(\mathbf{r})$. In this case the representation in the spin orbital basis is

$$g^{\Omega}(\mathbf{x}) = \sum_{ij} \psi_i(\mathbf{x}) G^{\Omega}_{ij} \psi_j(\mathbf{x}) \tag{3–35}$$

where

$$G^{\Omega}_{ij} = \sum_{ij} S^{\Omega}_{ij}\delta(\sigma_i,\sigma_j) \tag{3–36}$$

and S^{Ω}_{ij} is the atomic overlap matrix (AOM) that results from restricting the integration to the domain Ω.

Diagonalization of G with a proper unitary matrix U yields a new set of domain natural (occupied) orbitals (DNO) $\boldsymbol{\eta} = \boldsymbol{\phi}U$ and their respective occupations n of $G^{\Omega}U = nU$ such that

$$g^{\Omega}(\mathbf{x}) = \sum_i n_i |\eta_i(\mathbf{x})|^2 \qquad N_{\Omega} = \sum_i n_i \qquad 0 \leqslant n_i \leqslant 1. \tag{3–37}$$

When the occupation n_i is close to 1 the associated DNO is localized in center A indicating that there is a lone pair or core electron. Other than that, the small values of n_i manifest the existence of delocalization over two or more centers. In such a case the DNO is a one electron density delocalized (with non-zero values in $\mathbf{r} \in \Omega^c$) outside the basin where we are integrating. Unfortunately for the analysis of multi-center bonding with DAFHs, the DNO has the symmetry of center Ω hardening the analysis of their spatial visualization and consecutive association with bonding types. Isopycnic transformations are useful but they are not an optimal solution. We will see how generalization of DAFHs allows a more convenient dissection of multi-center bonds.

The density of a monodeterminantal wavefunction can be recast into

$$\rho(\mathbf{r}) = \sum_i |\eta_i(\mathbf{x})|^2. \tag{3–38}$$

Remembering that we are partitioning the space in two regions $\Omega \cup \Omega^c = \mathbb{R}^3$, the following relations are satisfied

$$n_i^{\Omega} = n_i^{\Omega^c} \qquad \phi_i^{\Omega} = \phi_i^{\Omega^c} \qquad s_i^{\Omega} = 1 - s_i^{\Omega^c} = n_i^{\Omega} = 1 - n_i^{\Omega^c} \tag{3–39}$$

with

$$s_i^{\Omega} = \int_{\Omega} |\eta_i^{\Omega}(\mathbf{r})|^2 d\mathbf{r} \qquad s_i^{\Omega^c} = \int_{\Omega^c} |\eta_i^{\Omega^c}(\mathbf{r})|^2 d\mathbf{r} \tag{3–40}$$

From these relations it is straightforward to derive the sum rule in equation (3–32)

$$\begin{aligned}
N_{\Omega} &= \sum_i (n_i^{\Omega} s_i^{\Omega} + n_i^{\Omega^c}(1 - s_i^{\Omega^c})) \\
&= \sum_i (n_i^{\Omega} s_i^{\Omega} + (1 - n_i^{\Omega})(1 - (1 - s_i^{\Omega})) \\
&= \sum_i (n_i^{\Omega} s_i^{\Omega} + (1 - n_i^{\Omega})(s_i^{\Omega})) \\
&= \sum_i (n_i^{\Omega} s_i^{\Omega} + (1 - n_i^{\Omega})(n_i^{\Omega})) \\
&= \sum_i (n_i^{\Omega} n_i^{\Omega} + (1 - n_i^{\Omega})(n_i^{\Omega})) \\
&= \sum_i (n_i^{\Omega})^2 + \sum_i (n_i^{\Omega} - (n_i^{\Omega})^2) \\
&= \sum_i n_i^{\Omega}
\end{aligned} \tag{3–41}$$

becomes the sum of DNO occupations as in equation (3–37). From the penultimate equation we can see that localization and delocalization (except for a multiplicative factor) indices are

$$\lambda^{\Omega} = \sum_i n_i^2, \qquad \delta^{\Omega\Omega^c} = \int_{\Omega^c} g^{\Omega}(\boldsymbol{x})\, d\boldsymbol{x} + \int_{\Omega} g^{\Omega^c}(\boldsymbol{x})\, d\boldsymbol{x} = 2\sum_i n_i(1-n_i) \tag{3–42}$$

Given that single determinant functions are invariant under unitary transformations the transformation of the spin-orbital basis to the DNO basis does not change anything. In this basis acquired by orthogonal projection to the domains Ω and Ω^c the overlap matrix $\widehat{S}_{ij}$ of equation (3–17) is diagonal in both regions with eigenvalues π_i or $1-\pi_i$ substituted here by n_i or $1-n_i$.

Multideterminant wavefunctions have a similar expression for (3–35)

$$g^{\Omega} = \sum_{ij}^{N_b} \psi_i(\boldsymbol{x}) G_{ij}^{\Omega} \psi_j(\boldsymbol{x}), \tag{3–43}$$

replacing G_{ij}^{Ω} by

$$G_{ij}^{\Omega} = \sum_{kl}^{N_b} C_{ijkl} S_{kl}^{\Omega} \delta(\sigma_i, \sigma_j), \tag{3–44}$$

with coefficients C_{ijkl} needed for the expansion of the second order exchange-correlation probability density

$$\lambda_2(\boldsymbol{x}_1, \boldsymbol{x}_2) \equiv \rho_{xc}(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{ijkl}^{N_b} C_{ijkl} \psi_i(\boldsymbol{x}_1)\psi_j(\boldsymbol{x}_2)\psi_k(\boldsymbol{x}_3)\psi_l(\boldsymbol{x}_4) \tag{3–45}$$

in terms of spin-orbitals. Diagonalizing in the basis of correlated DNOs we obtain

$$g^{\Omega}(\boldsymbol{x}) = \sum_{i}^{N_b} n_i^{\Omega} |\eta_i^{\Omega}(\boldsymbol{x})|^2. \tag{3–46}$$

3.2.3 Natural Adaptive Orbitals

To end this section, and returning to the generalization of DAFH to greater number of centers we will consider an exhaustive partition of physical space $\mathbb{R}^3 = \cup_j^m \Omega_j$ such that integration of the coordinates of one electron over all space can be divided as $\int = \sum_j^m \int_{\Omega_j}$. Based on the equation (1–72) the electron population is recovered summing all condensations

$$N = \int \lambda_N(\mathbf{r}_{i\leqslant N})\, d\mathbf{r}_{i\leqslant N} = \prod_i^N \sum_j^m \int_{\Omega_j} \lambda_N(\mathbf{r}_1, \cdots, \mathbf{r}_N)\, d\mathbf{r}_i. \tag{3–47}$$

The average population of one region A is the same previous equation restricting the integration of one of the electrons to only the domain A. The same way, the average population of two centers is obtained with the restriction of two electron integrations to those domains. Furthermore, from the norm-conserving property of the recursive formula (1–71), we know that the smallest order cumulant needed to get the desired ν-center average population is λ_ν

$$N_{ij\ldots\nu} = \int_{\Omega_i}\int_{\Omega_j}\ldots\int_{\Omega_\nu}\underbrace{\int\cdots\int}_{N-\nu}\lambda_N(\mathbf{r}_1,\cdots,\mathbf{r}_N) = \int_{\Omega_i}\int_{\Omega_j}\ldots\int_{\Omega_\nu}\lambda_\nu(\mathbf{r}_1,\cdots,\mathbf{r}_N). \tag{3–48}$$

Given that the interchange of integrals does not affect the result it is possible to obtain densities averaged over an arbitrary number of centers $\rho_{ij\ldots\nu}(\mathbf{r})$ simply avoiding one of the integrations that covers all space

$$N_{ij\ldots\nu} = \int \rho_{ij\ldots\nu}(\mathbf{r})\,d\mathbf{r} = \int\int_{\Omega_i}\int_{\Omega_j}\ldots\int_{\Omega_\nu}\lambda_{\nu+1}(\mathbf{r}_1,\cdots,\mathbf{r}_N). \tag{3–49}$$

The most important ones are one-center averaged densities ($\rho_i = g_i^\Omega$) and two-center (A and B) averaged densities. The last requires the third order cumulant density

$$\rho_{ij}(\mathbf{r}) = \int_{\Omega_i} d\mathbf{r}_1 \int_{\Omega_j}\lambda_3(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)\,d\mathbf{r}_2. \tag{3–50}$$

That partition recovers the total density

$$\rho(\mathbf{r}) \equiv \lambda_1(\mathbf{r}) = \sum_{i,j}\rho_{ij}(\mathbf{r}), \tag{3–51}$$

when summing over all pairs of domains in the same fashion as DAFHs (3–34), and can be generalized to a maximum condensation of $N-1$ electrons in any $N-1$ centers (may be it the same region repeated $N-1$ times)

$$\rho(\mathbf{r}) = \sum_{i,j,\ldots,N-1}\rho_{ij\ldots N-1}(\mathbf{r}). \tag{3–52}$$

ν-center generalized densities can be represented in the basis of molecular orbitals

$$\rho_{ij\ldots\nu}(\mathbf{r}) = \boldsymbol{\phi}(\mathbf{r})D^{ij\cdots\nu}\boldsymbol{\phi}^\dagger(\mathbf{r}), \tag{3–53}$$

where $D^{ij\cdots\nu}$ is a symmetric matrix and $\boldsymbol{\phi}(\mathbf{r})$ is a vector containing all occupied orbitals. The result of diagonalizing $D^{ij\cdots\nu}$ is a set of natural adaptive orbitals (NAdOs: $\eta_{ij\ldots\nu}$) with eigenvalues n_a called natural adaptive occupations (NAdOccs)

$$\rho_{ij\ldots\nu}(\mathbf{r}) = \sum_a n_a^{ij\cdots\nu}|\eta_{ij\ldots\nu}(\mathbf{r})|^2. \tag{3–54}$$

The orthonormality of NAdOs $\int |\eta_{ij\ldots\nu}(\mathbf{r})|^2 = 1$ assures that the sum of all $n_a^{ij\ldots\nu}$ NAdOccs is equivalent to the Ω_i–Ω_j–Ω_ν delocalization index

$$N_{ij\ldots\nu} = \sum_a n_a^{ij\ldots\nu}. \tag{3–55}$$

And the case of $\nu = 1$, (3–41), is well known from the previous subsection.

In contrast to DAFHs, NAdOccs directly give the delocalization index

$$\delta^{\Omega_i\Omega_j} = N_{ij} + N_{ji} = \sum_a (n_a^{ij} + n_a^{ji}) \tag{3–56}$$

without requiring a two basin partition, that otherwise would reduce to $n_a^{ij} = n_a^i n_a^j = n_a^{ji}$ recovering (3–48). A sign that there are bonding contributions involving more than two centers is the spread of $\eta_{ij\ldots\nu}(\mathbf{r})$ outside of the regions Ω_i and Ω_j. Multi-center bonding indices can be captured if we integrate higher order cumulants into three-, four-,..., ν-center NAdOs until they are fully localized. NAdOs can be seen as a generalization of Natural Bond Orbitals (NBO) [158], in the same spirit promoted by Zubarev et al. [192], although this time correlation is included explicitly.

3.3 PARTITIONING REAL SPACE INTO CHEMICALLY MEANINGFUL ENTITIES

EDFs, DAFHs, and NadOs are tools that are settled to work for a broad spectrum of domains. Depending on the chemical property desired to be explained one or another domain partition may be better suited but if one of them enjoys universal acceptance is the atomic partition. The concept of an atom is very clear when we only consider it isolated. There, a set of atomic orbitals with boundaries at infinite describe the electron distribution when they are squared. But it vanishes as we embed it in a molecule. Atomic orbitals of different atoms are combined (LCAO) and no single solution exists to separate individual atomic terms.

Probably the first attempts to divide space in chemically relevant regions are those rooted in EDF distributions. Aslangul et al. proposed [6] a variational minimization of the missing information

$$I[P, \Omega] = \left[-\sum_{n=0}^{N} p_n \ln p_n \right] \tag{3–57}$$

where stationary Ω *loges* minimize the entropy of the distribution. In other words they tried to find the regions where the probabilities deviate more from a uniform distribution $p_n(\Omega) = 1/N$, without being strictly localized. Thus, in each loge there is a high probability of finding some number of electrons. This

idea went rather unnoticed and has been followed only recently by Savin [165], who searched for those domains that maximize the probability of a single event.

At a time that nobody payed attention to the electron density, Richard Bader et al. discovered [12] from close inspection of density maps that the similarity among different molecules of electron densities in regions that surround atomic nuclei is related to the transferability of kinetic energy densities. Immediately after, Bader and coworkers found that zero-flux surfaces of the density create a natural partition of space in mononuclear basins that maximizes the transferability of kinetic energy densities whatever definition of kinetic energy density is used. Based on these findings, an atomic virial theorem for stationary states was conjectured. Eventually, the virial theorem was proved [177, 178] so a solid foundation for atoms in molecules was settled.

The atomic virial and related atomic theorems that were later derived have in common a variational formulation of a free boundary problem that replaces the functional from which the Schrödinger equation is derived for the functional

$$I[\Psi, \boldsymbol{\nabla}\Psi, \Omega] = \int_\Omega d\mathbf{r}_1 \int F(\Psi, \boldsymbol{\nabla}\Psi)\, d\mathbf{r}'_1 \qquad F(\Psi, \boldsymbol{\nabla}\Psi) = \frac{1}{2}\sum_i^N |\boldsymbol{\nabla}_i\Psi|^2 + (V-\varepsilon)|\Psi|^2, \tag{3–58}$$

that integrates one electron over the variable domain and the rest $d\mathbf{r}'_1 = d\mathbf{r}_2 \cdots d\mathbf{r}_N$ over all space. Variations of the wavefunction $\delta\Psi = \Phi - \Psi$ originate three sources of modification in I:

$$\delta I[\Psi, \boldsymbol{\nabla}\Psi, \Omega] = \int_\Omega d\mathbf{r}_1 \int \frac{\delta F}{\delta\Psi}\delta\Psi\, d\mathbf{r}'_1 \tag{3–59a}$$

$$+ \oint_{\partial\Omega} d\mathbf{S}_1 \int \frac{\delta F}{\delta\boldsymbol{\nabla}\Psi}\delta\Psi\, d\mathbf{r}'_1 \tag{3–59b}$$

$$+ \oint_{\partial\Omega} d\mathbf{S}_1 \int F\delta\mathbf{S}_1 + \text{c.c.}. \tag{3–59c}$$

Both F (3–59a) and ΔF (3–59b) change, and within this type of variational problem the surface $\partial\Omega$ is also allowed to vary (3–59c). The reader interested in more details is referred to [11, 106, 130]. After some mathematical manipulations one arrives at the following condition for the stationary solutions

$$\delta\left[\int_\Omega \nabla^2\rho(\mathbf{r})\, d\mathbf{r}\right] = 0. \tag{3–60}$$

A sufficient condition for the solutions of this variational problem is that the stationary surfaces have zero net flux

$$\nabla\rho(\mathbf{r})\cdot\mathbf{n} = 0 \qquad \forall\mathbf{r}\in\partial\Omega \tag{3–61}$$

and that they do not go through the nuclei [42]. This separatrix divides the space in regions with different flow behaviour. We now briefly present the link with well known ideas from topology.

3.3.1 Topology induced by a vector field

In the same fashion that a massive particle blends space-time, the charge of an atom nucleus blends the electron density. At the position of the nucleus the curvature is maximum and as we go further away the curvature vanishes in the limit of infinite distance as it approaches zero. Since the forces acting among the protons and electrons are attractive, electrons will preferentially be placed near the nucleus hence the density will have a cusp. If we include another atom there will be two cusps. Starting from the position of one atom and going along the line that connects two atoms, the curvature of the density will be negative, as it decreases, but at some point it will have to turn its trend, a saddle point, and start to increase (positive curvature) to reach the next maximum located at the other nucleus. The ridge of maximum electron density concentration, in the directions perpendicular the path connecting the nuclei, is a trail of preferable electron interchange.

More than the electron density itself, it is the gradient field which is unveiling the characteristics of the density, standing out points where the gradient vanishes. These points, so-called critical points, are points with vanishing first derivatives ($\boldsymbol{\nabla}\rho(\mathbf{r}) = \mathbf{0}$). There are several types of critical points that can be distinguished using the Laplacian. The procedure would be the same for any field although the physical interpretation varies. For now, the density will be our guiding example. At a point of $\mathbb{R}^3$ space the electron density Laplacian is,

$$\nabla^2\rho = \frac{\partial^2\rho}{\partial x'^2} + \frac{\partial^2\rho}{\partial y'^2} + \frac{\partial^2\rho}{\partial z'^2} = \mathrm{Tr}(\mathbb{H}_\rho) = \sum_{i=1}^{3} \lambda_i, \tag{3–62}$$

the sum of diagonal components resulting from the diagonalization of the Hessian ($\mathbb{H}_\rho$) in principal curvatures (λ_i). Provided that the scalar field is smooth and has no degenerate critical points (non-singular Hessian) nor critical points with the same value (image), then critical points compress the global shape of the density.

Critical points are characterized by the rank of the Hessian (r) and signature ($s = \sum_i \lambda_i/|\lambda_i|$), commonly abbreviated as (r, s). There are four types of non-degenerate density critical points:

- Nuclear critical points $(3, -3)$: A maximum in the three main axes. They are almost always located at nuclei positions.
- Bond critical points $(3, -1)$: they are located in a minimum between two nuclei. The density is a maximum in the other two directions. They are located at the border of two basins. The set of points connecting it with other nuclear points following the maximum gradient define what has been called *bond paths*.
- Ring critical points $(3, +1)$: They appear in the center of aromatic rings.
- Cage critical points $(3, +3)$: A minimum in the tree main axes.

Bond critical points are the most interesting ones because the properties at them can be correlated with chemical behavior. The density value is related to the strength of covalent bonds. The Laplacian values are related to the strength of non-covalent (closed shell) interactions. Only equivalent bond pairs can be compared. The ellipticity is often used as an indicator of π character for an interaction. The ellipticity, using $\lambda_1 < \lambda_2 < \lambda_3$, of a bond critical point is defined as

$$\epsilon = \frac{\lambda_1}{\lambda_2} - 1. \tag{3–63}$$

It gives a measure of the electron distribution anisotropy in the principal axes perpendicular to the bond path. For σ and triple bonds cylindrical symmetry is presented and the ellipticity value is ≈ 0, for π interactions electron concentrates preferentially in one of the planes which give a maximum for the charge at the critical point, so its value is $\neq 0$.

The topology of any scalar field is determined by its critical points and asymptotic behavior at an infinite sphere, the mathematical relation is contained in the Hopf-Poincaré relationship. In molecules the relationship of the electron density is

$$n - b + r - c = 1, \tag{3–64}$$

where n, b, r and c denote the number of nuclear, bond, ring and cage critical points. In periodic systems the asymptotic behavior is different, the relation equals 0.

The content of this section could be called chemical topography. The following section describes how to assign physical properties to chemical entities using the topology of the field.

3.3.2 The atomic partitioning of molecular properties

Topological methods aim at the segmentation of a vector field into real space areas of different flow behavior. All vector lines can end either at an attractor (nuclear critical point) or at a saddle point and the spatial region given by all points whose vector lines end at the same attractor are called basins. In the dynamical systems jargon they are stable manifolds. **QTAIM** selects attractor basins to divide the real space domain. The surface that delimits the basins obeys the zero-flux condition

$$\frac{\partial \rho(\mathbf{r})}{\partial x}\frac{\partial S(\mathbf{r})}{\partial x} + \frac{\partial \rho(\mathbf{r})}{\partial y}\frac{\partial S(\mathbf{r})}{\partial y} + \frac{\partial \rho(\mathbf{r})}{\partial z}\frac{\partial S(\mathbf{r})}{\partial z} = 0. \tag{3–65}$$

This topological partition distinguishes from alternative partitions for being exhaustive and in real space. That is, basins recover the whole space, $\mathbb{R}^3 = \cup_A \Omega_A$, while being mutually exclusive, $\Omega_A \cap \Omega_B = \emptyset$.

For domains defined by a separatrix with the above property, the average potential is the average of the virial forces applied to them. As well, the hyper-

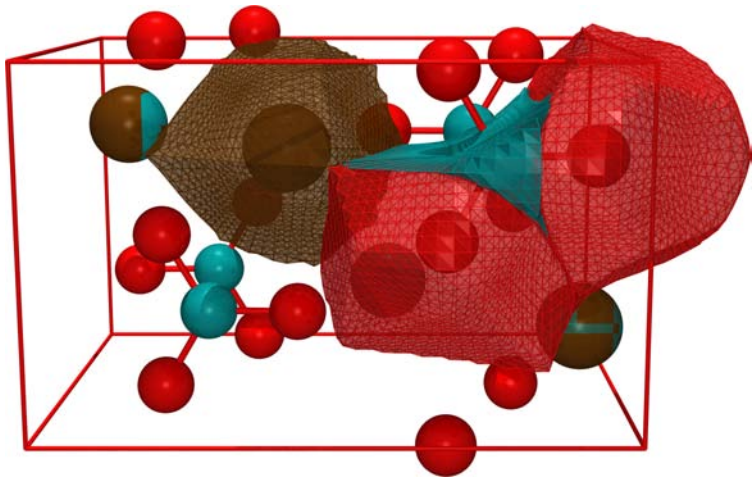


Figure 3: Aragonite structure and its non equivalent atomic basins: Ca (green), C (brown), O (red).

virial theorem leads to the definition of domain properties after the integration of any mono-electronic quantum operator (O)

$$O_\Omega = \frac{N}{2}\int_\Omega d\mathbf{r}\int [\Psi^* O\Psi + (O\Psi)^*\Psi]\; d\mathbf{r}'. \tag{3–66}$$

Note that the kinetic energy values are invariant when the zero-flux condition is met. Having a measure of the local potential and a unique definition of local kinetic energy we have definitively a decomposition of the energy in domains. Unifying all the clues disposed might guess us that the basins can be identified as proper quantum atoms.

Two of the most important properties are the volume, V_Ω, and electronic charge, N_Ω, of each molecular subunit Ω

$$V_{\Omega,\rho} = \int_\Omega d\mathbf{r}, \qquad N_{\Omega,\rho} = \int_\Omega \rho(\mathbf{r})\, d\mathbf{r}.$$

Since the properties are additive,

$$\langle O\rangle = \sum_\Omega O_\Omega, \tag{3–67}$$

Ω may be a single atom or an atomic group. The net charge of Ω is the sum of nuclear and electronic charge, $Q_{\Omega,\rho} = Z_\Omega + N_\Omega$. These formulas will be useful to figure out intermolecular charge transfer in Chapter 7. Moreover, the integration could be over basins of other fields.

3.3.3 The electron density Laplacian

The Laplacian is a measure of the density field curvature. When the density has a maximum the curvature is negative, $\nabla^2\rho(\mathbf{r}) < 0$, and when it is a minimum the curvature is positive, $\nabla^2\rho(\mathbf{r}) > 0$. Thus, the Laplacian identifies local charge depletions or concentrations.

The values of the Laplacian for an atom oscillate along radial directions, between positive and negative values. The function reveals the atomic shell structure. Each shell has a local maxima and minima along a radial direction. The outer shell is the most important because it is involved in bonding interactions. If it is a charge accumulation shell we call it a Valence Shell Charge Concentration (VSCC), else if it is a charge depletion we call it Valence Shell Charge Depletion (VSCD). Local maximum points $(3,-3)$ and local minima points $(3,+1)$ in VSCC regions are physically interpreted as nucleophilic and electrophilic sites in the acid-base Lewis theory. Critical points are also applied to identify the lone pairs of Valence Shell Electron Pairs Repulsion (VSEPR) theory.

The main characteristic of a covalent bond is that the valence shell electrons are shared. The value of the Laplacian at the bond critical point is negative. Non covalent interactions (e.g. chalcogen, halogen, hydrogen, ionic) involve two closed shell atoms. The Laplacian is positive at bond critical points and is related to the strength of the interaction.

The results obtained from integration over topological basins are negative [105, 136].

In the next section we will focus on the Electrostatic potential.

Other fuzzy partitions exist where a point in real space is not unequivocally assigned to an atom. Develop the weight functions

3.3.4 The electrostatic potential and its meaning

The Molecular Electrostatic Potential (MEP) at a point $(\mathbf{r})$ in space $(\mathbb{R}^3)$ is defined as the energy required to bring a positive particle of unit charge from the infinite to that point. In a molecule the potential at infinite is zero by definition, $\phi_0 = \lim_{\mathbf{r}\to\infty}\phi := 0$. It is also the interaction energy between the net charge density of a molecule, $\rho_t(\mathbf{r})$, and a test positive unit particle that mimics a proton, their interaction energy being evaluated without polarization or nuclear rearrangement effects,

$$\phi(\mathbf{r}) := \int_V \frac{\rho_t(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' + \phi_0. \tag{3–68}$$

The net charge contains the contributions of the electron density distribution and the nuclei $\rho_t(\mathbf{r}) = \sum_A Z_A\delta(\mathbf{r}-\mathbf{R}_A) - \rho(\mathbf{r})$. It is important to note that the Electrostatic Potential is a physical observable that can be determined by X–Ray Diffraction.

The seminal paper [24] by Bonaccorsi, Scrocco and Tomasi in 1970 brought attention to the Molecular Electrostatic Potential (MEP) in quantum chemistry

for the first time. Since then, many works have emphasized the role of MEP determining the interactions of a molecule with its environment [129]. The interaction between two molecules placed at long distances one apart another is dominated by the electrostatic potential with leading $1/r$ terms.

The study of the topology of the MEP was started by Gadre *et al.* [141]. Tsirelson identified zero-flux surfaces in solids [187] and Mata and co-workers investigated the zero-flux surfaces and the alternative nucleophilic/electrophilic partitioning [113] using the multipolar model.

Our approach to obtain the Electrostatic Potential (ESP) in a crystal is the same. The electrostatic potential is related to the electronic density by the Poisson equation

$$\nabla^2\phi(\mathbf{r}) = -4\pi\rho_t(\mathbf{r}). \tag{3–69}$$

Solving for ϕ we obtain equation 3–68, where nuclei positions are given by structural determination and the electron density by the multipolar model. We cannot recover the potential of reference (ϕ_0), so it is arbitrary. The absolute values of the ESP are arbitrary too, otherwise, the relative values are kept intact.

Properties of the scalar field

The ESP field exhibits the same type of critical points found in the electron density: maxima, minima and saddle points are the most frequent. Thus, the ESP can be characterized by critical points of the gradient field, which is minus the electric field, $\mathbf{E} = -\boldsymbol{\nabla}\phi(\mathbf{r})$. [1] The Hessian $\mathbb{H}_\phi$ is determined by the curvatures of the field γ_i

$$\mathbf{P}\mathbb{H}_\phi\mathbf{P}^{-1} = \begin{pmatrix} \gamma_1 & 0 & 0 \\ 0 & \gamma_2 & 0 \\ 0 & 0 & \gamma_3 \end{pmatrix}. \tag{3–70}$$

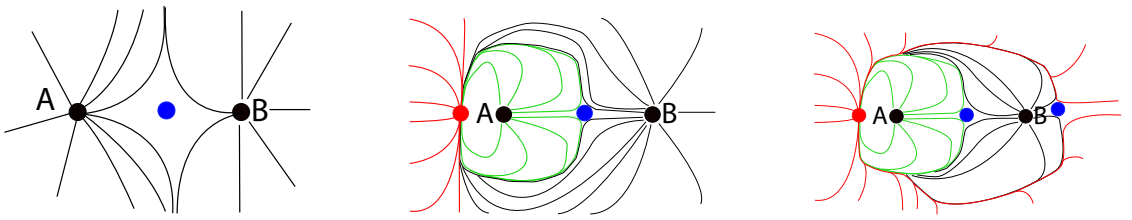
In contrast with the electron density, ESP maxima $(3,-3)$ can only occur at nuclear positions. Minima $(3,+3)$ identify local concentrations of electron density, such as lone pairs. Maxima and minima are identified as nucleophilic and electrophilic sites. $(3,-1)$ saddle points are analogous [58] to density bond critical points.

In contrast with the electron density, the ESP has positive (electrophilic regions) and negative values (nucleophlic regions). For example, in a cation (see figure 4), gradient field lines go from a positive value at nuclei to zero at infinite. In an anion, gradient field lines go from zero at infinite to a negative value at a localization of electron charge density. In a neutral molecule, gradient lines go from a positive value at nuclei to infinite and return to end at a minima describing a loop that fills all the space. Therefore, the asymptotic behavior is different and so is the Poincaré–Hopf relationship:

$$n_{-3} - n_{-1} + n_{+1} - n_{+3} = n_+ - n_-. \tag{3–71}$$

1 The zero potential has no effect in the topology.

Cation: A^+—B^+ Neutral: A^-—B^+ Anion: A^-—B



$2-1+0-0=1-0$ $2-1+0-1=1-1$ $2-2+0-1=0-1$

Figure 4: Diagrams of projected electric field lines for hypothetic AB diatomics and the Poincaré-Hopf relationship (3–71). Zero-flux surfaces have no gradient line crossing them. Critical points appear where gradient lines diverge or converge. Out of the plane critical points also appear, making the interpretation more difficult.

The terms in the left hand side are the number of critical points with signature s (n_s). The terms in the right hand side are the number of asymptotic maxima (n_+) and minima (n_-).

The Laplacian at a EBCP [2] (figure 5) is related to the total density at that point by the Poisson equation, (3–72). We can relate the strength of an electrostatic interaction with the value of the Laplacian.

$$\nabla^2\phi = \sum_i^3 \gamma_i = -4\pi\rho_t(\mathbf{r}) = -\nabla\mathbf{E} \tag{3–72}$$

We can even decompose [3] the electrostatic force in two components, an attractor term along the bond path and a repulsion term perpendicular to the bond using the curvatures of the field [115]

$$\mathbf{F}_E(\mathbf{r}) = -\rho(\mathbf{r})\mathbf{E} \approx \rho(\mathbf{r})\left[\gamma_{\|}(z-z_0)\hat{\mathbf{z}} + \gamma_\perp r\hat{\mathbf{r}}\right].$$

Orientation of the interactions rationalized with the aid of σ-holes

The electrostatic potential is positive almost everywhere. Negative regions can be clearly identified as nucleophilic regions that are expected to interact with an electrophile. Electrophilic regions are more difficult to find out because there is no well localized region with positive values. The solution goes through mapping the electrostatic potential on a representative low density isosurface such as the van der Waals surface.

If we map the ESP on a low density isosurface and search for maxima on it we find areas that act as lewis acids and can interact with charge donors . It

2 EBCP: $(3,-1)$ electrostatic bond critical point.
3 With cylindrical symmetry.

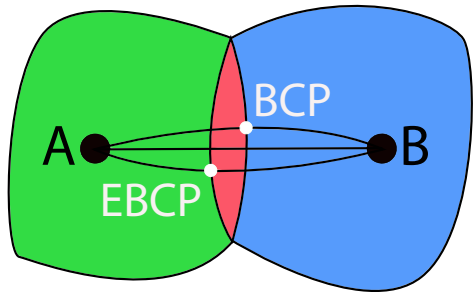


Figure 5: Topology of the electron density and electrostatic potential superimposed. Regions with partial negative (red), neutral (green), and positive (blue) charge conform the basins. Electron density basins of A (green+red) and B (blue) contain their respective atomic charges. ESP basins of A (green) and B (red + blue) are neutral. EBCP appears in the ESP surface. BCP appears in the density basins surface.

has been found that in atoms of groups 14 to 17 areas of low density appear on the opposite side of established sigma bonds (see figure 6), called σ-holes, that let the atoms act as acids.

A σ-hole is a low electron density, and therefore relatively positively charged, area adjacent to a σ-bonded element from groups 14 to 17. These areas appear on the opposite side of such atoms to the bond, along the same axis. The positive charges let the atoms act as Lewis acids, and bond non-covalently with electron donors. In this way σ-holes are located where charge densities provide viable bonding sites between atoms.

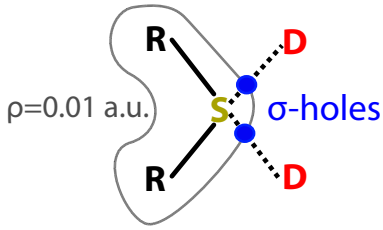


Figure 6: Sigma holes are maxima of the ESP on a low density isosurface.

Partition of the electrostatic field

Surfaces without any gradient line crossing them define the ESP basins (Λ) [187]

$$\nabla\phi(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = -\mathbf{E}(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \qquad \forall \mathbf{r} \in S(\mathbf{r}). \tag{3–73}$$

In the case of anions there is a region of space containing the negative charge of the molecule that does not belong to any atom. Otherwise closed basins have

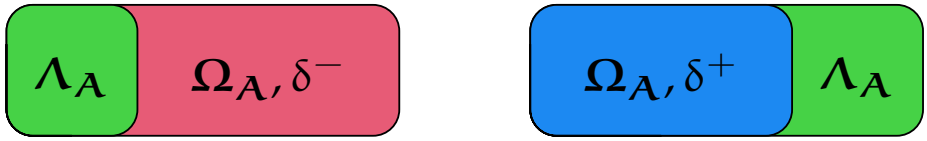


Figure 7: Basins of the electrostatic potential and the electron density. Anion (left), cation (right).

no net charge inside [113] because there is no electric field line crossing their boundary. Besides the net charge inside is 0 applying the Gauss theorem

$$Q_{A,\phi} = \int_{\Lambda_A} \rho_t(\mathbf{r})\, d\mathbf{r} = \int_{S_\phi} \mathbf{E}\cdot\mathbf{n}\, dS := 0. \tag{3–74}$$

ESP basins have no monopole interactions with the rest of the molecule so the calculation of interactions between basins should converge very fast. It is not clear if other multipoles should vanish but a sufficient condition to make all of them go to zero is exemplified by spherical symmetric distribution of charge in the basin (Newtons'theorem).

In anions the electron density basins (Λ) enclose ESP basins (Ω). The excess of charge falls outside ESP basins and inside electron density basins, figure 7. The opposite happens for cations. If we compare electron density and ESP basins we can find out where are partial charges located. What is still unknown is if those partial charges can be related to lone pairs or bonds.

At the EBCP the nuclear charge is fully shielded by the electronic distribution. This suggests that the value of the electrostatic potential at the EBCP is related to the strength of the ionic interaction between both atoms. [4] The position of the EBCP can also be related to the electronegativity of the interacting pair of atoms. If $(3,-1)$ points are present in both fields, the relative distance from an atom to the saddle point $d_{A-BCP} > d_{A-EBCP}$ implies that A is more electronegative than B, $\chi_A > \chi_B$.

3.4 PARTITION OF THE BINDING ENERGY

We have seen that within QTAIM the atomic energy partitioning is done transforming the Hamiltonian operator into an effective one electron operator. Then, the potential energy is obtained from the virial of the forces. Therefore, the computation of $\langle E\rangle_\Omega$ follows from the virial theorem. But out of equilibrium it does not hold so the present challenge is to cast the exact energy expression in a form that allows a resolution into physically interpretable expressions and to find out the principles that drive their interactions. Ultimately that information belongs to the wavefunction.

4 The potential at the EBCP is the Coulombic energy generated by the excess of charge (red region in figure 5) of A on the other atom (B in figure 5)

Attempts to get exact energetic partitions based on the wavefunction are futile because a separation of the wavefunction as a product of subsystem products $\Psi = \Psi_{\Omega_1}\Psi_{\Omega_2}$ is not possible. Nor electrons neither basis functions can be disentangled [127]. Although this approach is worth in its own, in fact it is a very good approximation in some cases, it is not what we are interested in. Since the energy is a magnitude that is carried by the second order density matrix we can think that a partition of this object would be enough and preferable.

3.4.1 Interacting Quantum Atoms (IQA) energy decomposition

In the following we focus our interest in QTAIM partitions defined by zero-flux separatrices to divide space in basins $\mathbb{R}^3 = \cup_{A=1}^{M}\Omega_A$. Following [106], the physical part of the functional $I[\Psi,\Omega]$ presented in equation (3–58) is

$$X_A = \int_{\Omega_A} d\mathbf{r}_1 \int \left[\frac{1}{2}\sum_i^N |\nabla_i\Psi|^2 + V|\Psi|^2\right] d\mathbf{r}'_1. \tag{3–75}$$

Recognizing the presence of only one- and two-body operators we can integrate out $N-2$ electrons and express it with the first order density matrix and the pair density (since the r_{12}^{-1} operator is diagonal)

$$X_A = \int_{\Omega_A}\left[\frac{1}{2}|\nabla\gamma_1(\mathbf{r};\mathbf{r}')|^2 + \sum_B^M \frac{Z_B}{r_B}\rho(\mathbf{r})\right] d\mathbf{r} + \frac{1}{2}\int_{\Omega_A} d\mathbf{r}_1 \int \frac{\rho_2(\mathbf{r}_1,\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2. \tag{3–76}$$

Besides, we may ask: how do we get a partition of the first order density matrix and pair density from a partition of the density $\rho(\mathbf{r}) = \sum_A \rho^A(\mathbf{r}) = \sum_A w_A(\mathbf{r})\rho(\mathbf{r})$ for any partition of unity

$$\sum_A w_A(\mathbf{r}) = 1 \qquad \forall\mathbf{r}\in\mathbb{R}^3? \tag{3–77}$$

Li and Parr showed [99] that a scheme addressing this question has to preserve indistinguishability of electrons and the physical meaning of the kinetic energy

$$\frac{T\gamma_1^A(\mathbf{r};\mathbf{r}')}{\gamma_1^A(\mathbf{r};\mathbf{r}')}\Big|_{\mathbf{r}'\to\mathbf{r}} = \frac{T\gamma_1^B(\mathbf{r};\mathbf{r}')}{\gamma_1^B(\mathbf{r};\mathbf{r}')}\Big|_{\mathbf{r}'\to\mathbf{r}} = \frac{T\gamma_1(\mathbf{r};\mathbf{r}')}{\gamma_1(\mathbf{r};\mathbf{r}')}\Big|_{\mathbf{r}'\to\mathbf{r}}. \tag{3–78}$$

as well as all other energy contributions, so that the partition is

$$\begin{aligned}\gamma_1(\mathbf{r};\mathbf{r}') &= \sum_A w_A(\mathbf{r}')\gamma_1(\mathbf{r};\mathbf{r}') = \sum_A \gamma_1^A(\mathbf{r};\mathbf{r}')\\ \rho_2(\mathbf{r}_1,\mathbf{r}_2) &= \sum_{A,B} w_A(\mathbf{r}_1)w_B(\mathbf{r}_2)\rho_2(\mathbf{r}_1,\mathbf{r}_2) = \sum_{A,B}\rho_2^{AB}(\mathbf{r}_1,\mathbf{r}_2),\end{aligned} \tag{3–79}$$

with the characteristic functions of each basin $w_A(\mathbf{r})$.

Moreover, X_A has a series of energetic terms that parallel those of the electronic energy for the entire molecule

$$X_A = T^A + V_{en}^{AA} + V_{ee}^{AA} + \sum_{B \neq A} \left(V_{en}^{AB} + \frac{1}{2} V_{ee}^{AB} \right), \tag{3–80}$$

where

$$\begin{aligned} T^A &= \frac{1}{2} \int_{\Omega_A} |\nabla \gamma_1(\mathbf{r}; \mathbf{r}')|^2 \, d\mathbf{r} \\ V_{en}^{AB} &= -Z_B \int_{\Omega_A} \frac{\rho(\mathbf{r})}{\mathbf{r}_B} \, d\mathbf{r} \\ V_{ee}^{AA} &= \frac{1}{2} \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_A} \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{\mathbf{r}_{12}} \, d\mathbf{r}_2 \\ V_{ee}^{AB} &= \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} \frac{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}{\mathbf{r}_{12}} \, d\mathbf{r}_2. \end{aligned} \tag{3–81}$$

A clarifying reorganization, inspired by McWeeny's ideas, is to agglomerate mono-centric and self-interaction terms as atomic net energies

$$E_{net}^A = T^A + V_{en}^{AA} + V_{ee}^{AA} \tag{3–82}$$

and, as well, two-body interatomic interactions

$$E_{int}^{AB} = V_{nn}^{AB} + V_{en}^{AB} + V_{en}^{BA} + V_{ee}^{AB}. \tag{3–83}$$

Therefore, the total molecular energy is

$$E_{mol} = \sum_A X_A = \sum_A \left[E_{net}^A + \sum_{B \neq A} E_{int}^{AB} \right]. \tag{3–84}$$

The large nucleus-nucleus, electron-nucleus and electron-electron quantities nearly cancel each other when they are collected to define E_{int}^{AB}. The tiny values yielded account for the subtle changes during chemical interactions. On the other hand, it does not happen the same for the net energy. Taking advantage of the fact that at dissociation the net energy reaches the energy of an isolated atom, the large intra-atomic components can be reduced subtracting the energy of the vacuum or any other reference state, to get a deformation energy

$$E_{def}^A = E_{net}^A - E_{ref}^A. \tag{3–85}$$

If the reference is the vacuum energy then the binding energy can be calculated as

$$E_{bind} = \sum_A E_{def}^A + \frac{1}{2} \sum_{A \neq B} E_{int}^{AB}. \tag{3–86}$$

Once the energy expression has been condensed to the maximum (3–86) one can envisage other more verbose ways to recast E_{int}, dividing V_{ee} in Hartree potential energy, exchange, and correlation, as in DFT

$$V_{ee} = V_H + V_X + V_C, \tag{3–87}$$

so that classic interaction energies

$$V_{classic} = V_{nn} + V_{ne} + V_{en} + V_H \tag{3–88}$$

can be separated from exchange and correlation V_{xc}, such that

$$E_{int} = V_{classic} + V_{xc}. \tag{3–89}$$

With this second rearrangement we achieve a cancellation of terms in $V_{classic}$. Including electrons and the nucleus altogether a large part of the interaction cancels, whereas the energy could be said to be due to small changes of the valence electronic distribution. So atoms without much charge transfer would have small classical interactions. Then the interaction is almost dominated by V_{xc}, which is a signature of covalent interactions.

To be more precise about the sources of deformation energy we can mention charge reorganization and charge transfer

$$E^A_{def} = E^A_{def}(CT) + E^A_{def}(CR), \tag{3–90}$$

where the former agglutinates the changes that affect to the electronic configuration of the neutral atom. The more familiar hybridization or promotion energies are part of that term.

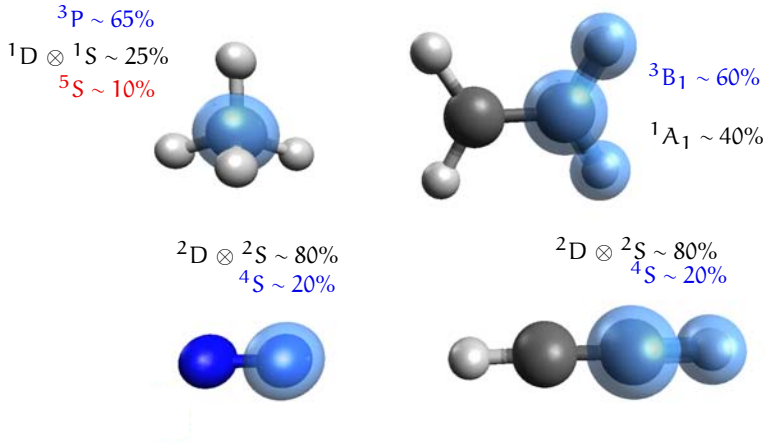
In essence, the virtue of this energetic decomposition is the ability to extract very detailed information of the interactions that can later be condensed in many ways to enlighten our perception of chemical interactions. Domains, as well, can be grouped together to describe the nature of, for example, functional groups.

Part II

How to appraise bonding energetics

4 TOWARDS A UNIQUE MEASURE OF BOND STRENGTH THROUGH INTRINSIC BOND ENERGIES

CONTENTS



Spin states weight for carbon in methane; methylene in ethene; nitrogen in the molecule N2; and carbine in acetylene (marked in blue). All species are considered at equilibrium.

Chemical substances are conglomerates that result from the constellation of atoms, some of them tightly linked together. This bonding of atoms becomes an essential part for the study of chemical compounds as it constitutes the building glue needed to stabilize more complex molecules. Moreover, the relation of structure with reactivity rests on the properties of chemical bonds. Often, synthetic routes achieve their goal of reaching a desired product by breaking the weakest bonds in the molecule, requiring us to know their relative strengths, to substitute some fragments by others. At this point we envisage the need of methods to calculate bond strength.

Despite being a very intuitive concept there is no physical observable associated to bonding yet. This also implies that there is no related quantum mechanical operator. But that does not take away from the fact that there is a reasonable chance that developing a measure of bond strength can make a significant contribution to the understanding of chemical reactivity. However, from the above description it appears that the energy is an essential aspect of bonding. While there are also other properties like bond length, density at the associated critical point, or stretching force constant, that have intimate connection with the bond strength, the energy seems the natural choice. For this reason we will focus on energetic measures of bond.

4.1 THE QUEST FOR RELIABLE BOND ENERGY DESCRIPTORS

Maybe the easiest and most direct approach to do measurements about the energetic stability contributed by a bond are thermodynamical experiments. There are two variables frequently used and tightly related to bonding. Both are taken from the process of bond making/breaking. The enthalpy change in the process

$$A - B(g) \rightarrow A(g) + B(g) \tag{4–1}$$

of fragmenting 1 mol of a molecule $A - B$ under standard conditions in two fragments A and B that were originally linked by a single bond shall be defined as the bond dissociation enthalpy (BDH). BDH quantities in turn reduce to bond dissociation energies (BDE) in the limit that the temperature is 0 K and neither reactives nor products have zero point vibration. As well, they can be obtained from numerical computation of the ground state energies of the molecules involved

$$BDE(A - B) = E^{gs}[A] + E^{gs}[B] - E^{gs}[AB]. \tag{4–2}$$

Then, BDE and BDH can be related if all vibrational plus rotational plus translational effects are known. For the rest of the present chapter we will be more interested on BDE.

Even if BDE are reported elsewhere as a proper descriptor of bond strength only for diatomic molecules the strength of a bond can be effectively measured

using bond dissociation energies (BDE) from calculation or bond dissociation enthalpies (BDH) from experiment. However, taking methane as an example, the successive breaking of bonds

$$CH_4(g) \rightarrow CH_3(g) + H(g) \qquad \Delta H_r(1\text{ atm, }298\text{ K}) = 104\text{ kcal/mol} \quad (4\text{–}3)$$
$$CH_3(g) \rightarrow CH_2(g) + H(g) \qquad \Delta H_r(1\text{ atm, }298\text{ K}) = 106\text{ kcal/mol} \quad (4\text{–}4)$$
$$CH_2(g) \rightarrow CH_1(g) + H(g) \qquad \Delta H_r(1\text{ atm, }298\text{ K}) = 106\text{ kcal/mol} \quad (4\text{–}5)$$
$$CH_1(g) \rightarrow C(g) + H(g) \qquad \Delta H_r(1\text{ atm, }298\text{ K}) = 81\text{ kcal/mol} \quad (4\text{–}6)$$

reveals bond dissociation enthalpies with uneven values. reveals that a different amount of energy is needed to break each of the four C-H bonds. Every time you break a hydrogen atom off the carbon, the environment of those left behind is altered. That variability of BDHs undermines their use as bond energy indicators as all C-H bonds in methane are equal by symmetry. To get proper bond energies that obey the equivalence of all bonds imposed by symmetry, the different BDEs are averaged to define a new thermodynamical bond energy (BE). That is the same as one fourth of the reaction enthalpy of the atomization process

$$CH_4(g) \rightarrow C(g) + 4H(g) \qquad \Delta H_r(1\text{ atm, }298\text{ K}) = AE. \quad (4\text{–}7)$$

Although this procedure is immediate, there exist many molecules for which the procedure cannot be replicated since they can have several non equivalent bonds or not all of them have the same connectivity with the central atom. The answer commonly used assumes that bond energies can be transferred between different molecules provided that they connect the same pair of atoms. Only then we can calculate the bond energies of all the bonds present in the molecule. It is not necessary to mention that this approximation, that only considers the topology of the bond, is a very rough one.

In effect, by reducing the molecule to a molecular graph a lot of information is lost. A possible refinement would include the equilibrium bond length as well as an indicator. This way information about the geometry also takes part in deciphering bond energies. However, it does not ensure a satisfactory result because Badger's rule, which states that a shorter bond is always stronger, is not always obeyed. The definitive way to solve the issue is to include also the overlap. As the reader might guess from the commentaries poured in previous chapters the overlap has a strong dependency in the method used to calculate the wavefunction. If anything we need to have clear is that to define a proper bond energy we have to combine geometrical with electronic structure information altogether.

The bond energy problem does not end there. Less obvious than the issue of transferability is the appropriateness of atomization energies. Recap that two basic assumptions were necessary to get bond energies: (i) the sum of all bond energies will be the same as the atomization energy given by a stepwise breaking of all bonds

$$AE = \sum_i BDE_i = \sum_i BE_i, \quad (4\text{–}8)$$

(ii) equal bonds have equal BEs, so they can be transferred. AE (BDE) are dynamical quantities that depend on the reactive as well as the products of the atomization (dissociation), they are a delta of deltas. Hence it cannot be said that AE are indicators of solely the resistance that the molecule opposes to break apart in atoms, even for a diatomic. But the stability of the resulting atoms also counts. This situation can be exemplified by several dramatic examples. $2NaH \rightarrow H_2 + Na_2$, $2HgO \rightarrow Hg_2 + O_2$, and $2HgCl \rightarrow Hg_2 + Cl_2$ are endothermic reactions and they only involve the breaking of one bond. Would this mean that the bond is destabilizing? The same could be said of many high-energy molecules. Or the near to 90 kcal/mol energetic difference between the hydrogenation of dinitrogen and acethylene suggest that the triple bond of dinitrogen is so much stronger than that of acethylene [25]? It seems more plausible to associate part of the energetics to the stabilization of the fragments. AE are reflecting the instability of a electronic configuration with respect to the dissociated states but not of the bond.

In the end we are interested in a bond energy descriptor that depends only in the stabilization generated by the specific characteristics of the bond and gets rid of the superfluous stabilization of the molecule relative to the fragments. For this purpose valence states, i.e. states of the fragments that retain the electronic structure exhibited in the molecule, are key. We could devise an in silico scheme to separate energetic contributions inside the atomization energy into *intrinsic bond energies* (IBE) [100, 101, 173] and relaxation effects [39] (E_R)

$$IAE = AE + E_R; \qquad IAE = \sum_i IBE_i \tag{4–9}$$

splitting the dissociation process in two steps: first scaling bond lengths with a frozen electronic configuration; then turning to their most stable isolated electronic states. This last component (E_R) would encompass promotion, hybridization, geometric freezing and charge reorganization terms. The adequacy of this argument has been extensively rationalized by Cremer et al. [39].

New types of bonding, such as tetrylones [55, 56, 185] or quadruple bonding in representative elements might now be accepted when those states rather than the fundamental ones represent the backing for bonding measurements.

4.1.1 Controversial views

However, when the atomic reference states are not clearly established bond strengths can be misleading [39]. This is the case of carbon atoms when they bond as C_2. Although it is an apparently easy to understand molecule, it has uncovered a lengthy discussion full of elaborate arguments to sustain or tear down the possibility of a quadruple bonding [41, 54, 70, 167, 168]. The origin of the dispute is not the multiplicity. All seem to agree in that there are four components of bonding with different spatial distributions, or so to speak electrons adopt four different strategies to reach maximum stability. Instead,

the confrontation is in the strength of the bond, on whether the bond of C_2 is stronger than in acethylene or not.

There are two arguments encountered in those articles that should not be used as the only evidence to support the choice of the reference state. One is that the reference of a n-bonded molecule is a state with n unpaired electrons. The idea behind it is that for creating a covalent bond two orbitals occupied by different spin electrons and localized in the two fragments would couple. The counterexample could be, by way of orbital reasoning, a dative bond. In that situation there would be no need to unpair electrons. Other argument is that the valence state of a fragment can be reached following the dissociation path that has no avoided crossings. The changes that are reflected on the potential energy curve or in the molecular state affect to the whole molecule and should not be taken as a justification for the election of valence state. Besides, only pure spin states could be considered. It is highly unlikely that the state of an atom-in-the-molecule will correspond exactly to a dissociated state.

Shaik et al. made an enthusiastic effort to understand the reference state with quasi-classical states or coefficients of either valence structure expansions or determinants made of molecular orbitals showing that the valence structures with more weight at equilibrium are related at dissociation to a pair of quintuplet states [169]. Whether those structures are capable of capturing the atomic states in-the-molecule and to show them clearly is a matter that should be addressed. The history of the C_2 bonding debacle and the rising number of controversies in computational chemistry [71] should not prevent us from trying to establish a firm ground for those concepts that are shaky.

Taking the states of the fragments out of the molecule is difficult in general. In fact, it requires from us to give a precise meaning to an atom in a molecule, be able to identify the contribution of each atomic state and simultaneously compute their energies (irrespective of the type of wavefunction). We aim to identify systematically the reference, in a way that the procedure to select valence state is always the same and does not depend on unjustified heuristic chemical insight.

4.2 INTRINSIC BOND ENERGIES IN THE REALM OF REAL SPACE THEORIES

Modern real space theories of bonding constitute an ideal framework to define atoms-in-the-molecule and examine their properties. For instance, QTAIM gives good definition of atoms that is invariant to changes of the method used to calculate the wavefunction whilst IQA separates energetic contributions in a convenient manner. Energetic terms devised from IQA (for a brief overview see section 3.4.1) partitioning comply with the separation of relaxation from interaction terms needed to compute intrinsic bond energies. Furthermore, the QTAIM restriction to equilibrium geometries imposed by the virial theorem is

no longer an issue. Electronic distributions are the perfect complement to IQA as the foundation for electronic spin states quantification. As if this were not enough, we will put an eye on the spatial distribution of bonds with orbitals adapted to include correlation or natural adaptive orbitals (NAdOs). At the end we hope to find a solid basis for further discussions.

4.2.1 Relaxation and interaction energies

iqa: intrinsic bond energies

qcs: in-situ bond energies(no polarization)

In the context of IQA the atomization energy is conveniently organized as the sum over all A atoms of their intra-atomic components E^A_{net} and their interactions $E^{A,B}_{int}$ with the rest of atoms B.

$$AE = \sum_A \left(E^A_{net} + \sum_{B>A} E^{A,B}_{int} \right). \tag{4–10}$$

Intrinsic bond energies find a very natural place in the inter-basin interaction energy whereas to have an equivalent to the relaxation energy we have to refer to the ground state energies of the fragments

$$E^A_{def} = E^A_{net} - E^A_{gs}. \tag{4–11}$$

E^A_{def} accounts for the charge reorganization that happens while the fragment is embedded in the molecule in the same way that Cremer describes for the relaxation energy. From now on relaxation energies will correspond to deformation energies (E^A_{def}) and intrinsic bond energies to interaction energies ($E^{A,B}_{int}$). You see that with very small changes we already have a way to measure intrinsic bond energies.

The most complicated part of our analysis regards the identification of atomic spin states and their linkage to the changes that E^A_{def} exhibit as will be seen in the next subsection.

4.2.2 Tracking local spin states

Although we are not in a position to assess spin state energetics with IQA alone, a combined effort with electron population distributions can clarify some questions. Before addressing the relation of spin states and EDFs we will make some general considerations: (i) molecules will be bi-partitioned $\Omega \cup \Omega^c = \mathbb{R}^3$, resonance structures denoted by (n_Ω, n_{Ω^c}), (ii) total wavefunctions will be restricted to be a simultaneous eigenbasis of the Hamiltonian and total spin operators

$$\hat{S}^2\Psi = S(S+1)\Psi, \qquad \hat{S}_z\Psi = \frac{1}{2}\sum_i^N (n_i^\alpha - n_i^\beta)\Psi = \frac{1}{2}(N^\alpha - N^\beta)\Psi = M_S\Psi. \tag{4–12}$$

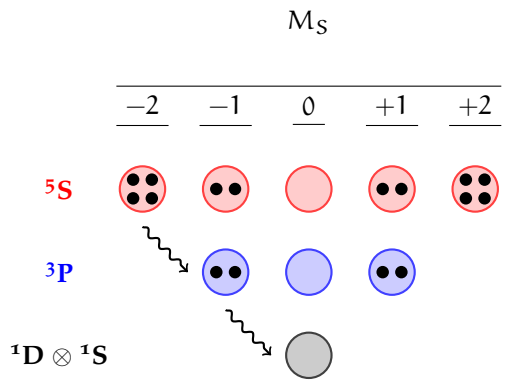


Figure 8: Electron distribution functions count electrons in space either as alpha or beta electrons. Resonance structures, here represented as columns, are labeled accordingly by the M_S eigenvalue. The structures with maximum multiplicity are the key to disentangle the states. The $M_S = \pm 2$ states are definitively part of a quintuplet (or higher spin) state, thus multiplying that probability by the degeneracy of the state (five in this case) we get the weight of the quintuplet. To maintain the normalization we have to subtract those from the pile with two excess electrons (twice) and no excess electrons (once), which is equivalent to removing the upper row. Applying this technique recursively to the remaining structures we end up with singlets, where no further separation can be done unless we have additional information from other sources. A dot represents an unpaired electron.

with average net spin quantum number zero ($S = 0$) and spin component $M_S \in \{-S, \cdots, S\}$ also zero — a singlet state. That is, the number of alpha and beta electrons in the whole system is equal ($N^\alpha = N^\beta$). Otherwise, this might not be true in each fragment.

Counting spin labeled electrons in both domains is not very different from plainly counting electrons. The formalism, that was first introduced in section 3.2.1, declares a probability space made of all partitions of the total number electrons in Ω and its complementary space, and alpha or beta labeling, the outcomes being Spin Resolved Resonance Structures [112] $S \equiv \{n^\alpha_\Omega, n^\beta_\Omega, n^\alpha_{\Omega^c}, n^\beta_{\Omega^c}\}$. For instance, α/β-population probabilities computed in the basin are already a measure of local spin [159], although not a complete one, since we have no well defined spin-squared operators within domains, $S^2_\Omega \Psi = S_\Omega(S_\Omega + 1)\Psi$.

During the last decade there has been a flurry of interest in local spin by several authors for its applicability to a better understanding of the interaction between open-shell subsystems. Several schemes have been proposed with a common origin in the work of Clark and Davidson [35]. There, they used a Hilbert-space projection of the spin operators in the basis of atomic orbitals as a sum of atomic and diatomic components. However, that method, and others derived upon it, have not uniquely defined projectors because they depend on the choice of the basis set. Another difficulty comes with correlated functions. Approaches inspired in population analyses take as fundamental variable the spin density matrix [1, 120, 186]. And yet the spin quantum numbers predicted with them are not always in accordance with some well known necessary conditions [81, 121]: (i) closed-shell RHF wavefunctions have local singlet states (ii) in the dissociation limit the values are the same as for the free atoms. (iii) in the limit of no correlation reduces to the determinantal values. Later works have been correcting those unphysical results. The first application of the local spin concept to real space is that of Alcoba et al. [2].

Returning to our alternative approach to get local spin, we take for granted that we know the probability that each spin resolved structure has. To proceed with the determination of local spin we take the maximum multiplicity $|M_S| = S$ structures first, those that have $2|M_S|$ unpaired electrons. For a schematic representation see figure 8. The weight of the highest spin state is straightforward. It is the probability of one of the structures $M_S = S$ or $M_S = -S$ multiplied by the multiplicity of the state $(2S + 1)$. Each atomic multiplet has M_S components that span from $-S$ to S. Besides, states of lower multiplicity within that spin-squared state have to be subtracted from the subsequent spin-squared state. Applying this recursive algorithm we reach an endpoint in the lowest spin-squared state. Remark that only when there is one spin state per spin quantum number S we can fully decompose spin states. In that situation we have to infer the spin states out of mixed probabilities by other means. The components of lower multiplets mix with those of greater spin. The probabilities are reported after the degeneracy of each atomic spin state has been accounted, from $-S$ to S. Our approach is correct in the dissociation limit and reduces to the single-determinant case when there is no correlation. In contrast

with the schemes presented up to now, in our treatment correlation is included without fuss. We have a set of limitations that are different, we believe that can contribute to a development of the local spin concept. Our insight comes from having information about the probability that each electron condensation has and not just their average. Summing up, the quest for a (coarse-grained) local, unique real space representation, of the S^2 operator, $S^2_\Omega \Psi = S_\Omega(S_\Omega + 1)\Psi$, to help us polish up the analysis, and therefore we would know the S spin quantum number of the domains is still open. As will be seen below these structures allow us to study the genuine spin states of each fragment giving us more insight about the electronic arrangements at stages previous to bond formation.

In principle there are four variables that drive the behaviour of the electronic distribution. Since the two domains cover all space the population is fixed to be N. Therefore the last variable is completely determined. The singlet state global condition allows us to simplify the expression for the spin resonance probabilities of two centers as

$$p(n^\alpha_\Omega, n^\beta_\Omega, n^\alpha_{\Omega^c}, n^\beta_{\Omega^c}) = p(n^\alpha_\Omega, n^\beta_\Omega, N/2 - n^\alpha_\Omega, N/2 - n^\beta_\Omega) = p(n^\alpha_\Omega, n^\beta_\Omega) \quad (4\text{–}13)$$

to focus solely on one fragment knowing that the conditional probability of having a given excess (or defect) of alpha electrons in the domain Ω^c is one when the number of electrons with alpha spin is equal to those with beta spin in domain Ω and otherwise is zero. The same applies for beta electrons.

On top of the assumptions made we group together probabilities with fixed Ω population $n^\alpha_\Omega + n^\beta_\Omega = n_\Omega$. Each group corresponds to a Pauling resonant structure $\{n_\Omega, n_{\Omega^c}\}$ [147]. In particular, the structure with the same number of electrons inside each basin as their isolated counterparts represents a system without charge transfer between two domains Ω and Ω^c, that means a purely covalent system. Pauling structure probabilities are thus branched in several spin resolved resonance probabilities

$$\begin{aligned} p(n_\Omega = n) &= \sum_{x=0}^{n} p\left(n^\alpha_\Omega = x, n^\beta_\Omega = n - x\right) \\ &= \sum_{x=0}^{n} p\left(n^\alpha_\Omega = x | n_\Omega = n\right) \times p(n_\Omega = n) \end{aligned} \quad (4\text{–}14)$$

conditioned to have the above mentioned restriction or partitioned with weights given by the conditional probability of having n electrons in Ω.

Previous works already pointed [110], in accordance with chemical insight, that the probabilities of Pauling resonant structures where sharing of core electrons is involved have very low values. With the appropriate choice of the valence electrons we can adjust very accurately the calculated probabilities. We therefore use a bonding model that takes apart core electrons, so they do not contribute neither to covalent bonding nor to charge transfer. Instead of using a model of centers and electrons, in the frame of spin structures it is due

to propose a model of electrons distributed in spin-orbitals. For example, a model of n electrons in n_o^α α spin-orbitals and n_o^β β spin-orbitals is denoted by $(n_o^\alpha, n_o^\beta, ne)$. Henceforth, a way to interpret $(n_\Omega^\alpha, n_\Omega^\beta, n_{\Omega^c}^\alpha, n_{\Omega^c}^\beta)$ structures is as the result of arranging n_Ω^α electrons in all α-spinorbitals (ψ_Ω^α) and n_Ω^β electrons in β-spinorbitals (ψ_Ω^β) of basin Ω, and $n_{\Omega^c}^\alpha$ electrons in α-spinorbitals $(\psi_{\Omega^c}^\alpha)$ and $n_{\Omega^c}^\beta$ electrons in β-spinorbitals $(\psi_{\Omega^c}^\beta)$ of basin Ω^c.

THE MODEL OF QUASI-INDEPENDENT SPIN PARTICLES. It is reasonable to think what would happen if electrons were totally independent. Thereupon the relative weight of spin resolved structures with respect to the fixed Pauling structure $\{n_\Omega = n, n_\Omega^c = N - n\}$ is

$$\begin{aligned} p\left(n_\Omega^\alpha = x | n_\Omega = n\right) &= \frac{p(n_\Omega^\alpha = x, n_\Omega^\beta = n - x)}{p(n_\Omega = n)} \\ &\approx \frac{p(n_\Omega^\alpha = x) \times p(n_\Omega^\beta = n - x)}{p(n_\Omega = n)} \\ &\approx \frac{\binom{n_o^\alpha}{x}\binom{n_o^\beta}{n-x}}{\binom{n_o^\alpha + n_o^\beta}{n}} = h(x; n, n_o^\alpha, n_o^\beta), \end{aligned} \tag{4–15}$$

where $0 \leqslant x \leqslant n$. The first approximation considers $\alpha - \beta$ population independence. The second approximation goes further stating that all electrons are independent. Since the tree variables n_Ω, n_Ω^α, and n_Ω^β follow binomial distributions that share the same parameter, the result is a hypergeometric probability mass function h characterized by three parameters: (i) the population of the basin n, (ii) the number of spin-up orbitals (n_o^α) and (iii) the same for spin-down orbitals (n_o^β). With them we can guess the probability values.

Previously it has been shown that spin-less probabilities at saturated delocalization stages are close to a binomial distribution [110] and some kind of aufbau principle appears for spin resolved probabilities [112] when we measure the relative proportion of a given spin structure over all structures with the same spatial distribution, i.e. same number of electrons in each basin. The principle found there is now quantitatively described by a hypergeometric distribution. To demonstrate this we will take an example of a basin with four spin-orbitals

and two electrons ($n_o^\alpha = 2, n_o^\beta = 2, 2e$). Applying the simplification mentioned in equation (4–15) the probabilities

$$p(n^\alpha = 0|n_\Omega = 2) = p(n^\alpha = 2|n_\Omega = 2) = \frac{\binom{2}{0}\binom{2}{2}}{\binom{4}{2}} = \frac{1}{6}, \tag{4–16}$$

$$p(n^\alpha = 1|n_\Omega = 2) = \frac{\binom{2}{1}\binom{2}{1}}{\binom{4}{2}} = \frac{2}{3}. \tag{4–17}$$

form a vector $\mathbf{p}_2 = (p(2,0), p(1,1), p(0,2)) = (1/6, 2/3, 1/6)$, which clearly favors intra-atomic pairing over a binomial distribution $\mathbf{p}_2 = (1/4, 1/2, 1/4)$.

The same effect applies to all Pauling resonant structures. On the basis that the total probability distribution is a convex combination of Pauling structures we can claim that the net effect is that electrons liken to be coupled inside domains. The total distribution set as a convex combination of Pauling structures also favors pairing. Thus pairing is showed to be favorable in the independent model. We conclude that as soon as there are two electrons in a basin a *spin quenching* effect is present.

We will see that valence CASSCF calculations determine probability values at equilibrium in close agreement with those predicted from the model. The correlation that that is included with the multiconfiguration is mainly static, whereas at equilibrium the correlation is mostly dynamic.

4.3 COMPUTATIONAL DETAILS

The processes that we will examine are the atomization of methane

$$CH_4(^1\Sigma_g) \rightarrow C(^3P) + 4H(^1S), \tag{4–18}$$

scaling all C-H bond lengths with the same factor whereas bond and dihedral angles are maintained, the dissociation of dinitrogen

$$N_2(^1\Sigma_g) \rightarrow 2N(^4P), \tag{4–19}$$

and the cleavage of C-C bonds of ethene and acetylene

$$C_2H_4(^1\Sigma_g) \rightarrow 2CH_2(^3B_1), \tag{4–20}$$

$$C_2H_2(^1\Sigma_g) \rightarrow 2CH(^2\Pi). \tag{4–21}$$

In the last two reactions hydrogen positions were optimized at each point. Acetylene is broken in two carbyne fragments following the minimum energy path, oriented in a trans configuration [172] to reduce the repulsion of head-to-head filled orbitals, instead of a collinear rapprochement.

Size-consistent calculations with full-valence CASSCF functions were performed with `GamessUS` [166]. A Def2-TZVPP one-electron basis is used for all systems except for the dinitrogen dissociation, where a larger basis that includes d functions (Def2-TZVPPD) was needed to achieve size-consistency.

All EDFs have been obtained with our `EDF` code [50] reading atomic (QTAIM) overlap integrals generated with `Promolden` [145]. For the integrations in QTAIM domains the radii of the beta spheres chosen is roughly half the distance from the respective nucleus to the nearest bond critical point in order to get accurate results both for the charges and the energy. Almost all systems presented difficulties with geometries in the repulsive region either because a group was ionized (carbon in methane) or because a non-nuclear maxima appeared. Natural Adaptive Orbitals (NAdOs) were obtained with another internal code: `denmat`.

For methane we want to test if the ^5S state of carbon is formed to the extent that is often claimed. However, it is known from long ago that the valence state is lower in energy, as a superposition that includes contributions of intra-configurational states (^1D, ^1S, ^3P) in an unknown proportion [100]. Nevertheless, it could never be proven because no theoretical information was available due to the nonexistence of a proper framework to treat it. As a consequence the explanation was later assumed to be true. Here we provide the weight of each electronic state of neutral CH_4 along the atomization.

Triply bonded nitrogens are interpreted as the result of coupling two ^4P states. By casual coincidence that is the ground state of the isolated nitrogen atom so from that argument nitrogen atoms do not need to be excited to a higher excited state to create three bonds. So is suggested by the absence of irregularities in the energy dissociation curve. Another question that we will try to answer is if the formation of σ- and π-type bonds occurs at separate stages.

The behaviour encountered in N_2 and CH_4 should be somehow mimicked by methyne (CH_2) and carbyne (CH) radicals during their dimerization. Methynes, as nitrogen atoms, do not need to be excited to reach their state-in-the-molecule. That electronic state is presumed to be a triplet state since two bonds need to be created. Like the nitrogen atom, the state of carbynes in the acetylene molecule is thought of as a quartet [54] but the ground state of carbyne is a doublet ($^2\Pi$). Hence the need of a previous excitation in the course of the dimerization. Equipped with our tools we will prove how accurate is this picture.

4.4 RESULTS AND DISCUSSION

4.4.1 Methane

There are few more enduring concepts in the mind of a chemist than the hybridization of carbon to enable covalent bonding to other four atoms. The bonding description of methane from its constituents assumes that one electron

of the carbon 2s orbital promotes to a 2p orbital and then all valence orbitals (2s+3×2p) mix through a unitary transformation to create four single occupied hybrid orbitals of type sp_3 with the adequate geometrical orientations to form bonds in the directions of the vertices of a tetrahedron. This reasoning is more based on heuristics than on a priori theoretical evidence. In this regard we try to find an explanation with the aid of electron distribution functions.

We studied the concerted atomization of the most stable energetic state, a singlet $^1A_1$ state.

If any configurational change happens it should be reflected in the IQA deformation energies. This term incorporates charge reorganization and the configurational changes we are interested in. The deformation energy of the carbon atom exhibits self reorganization during the process of bond formation. A selected region close to the equilibrium (figure 9) suggests an internal change of carbon atoms. The energetic differences are very small to be associated with a $^3P \rightarrow {}^5S$ transition, but it is evident that there is an energetic barrier originated by a non-optimal configuration of the fragments for creating bonds. After reorganizing the system is stabilized. Two possible reasons may drive this behavior: charge reorganization or configurational changes. We seek to achieve a further decomposition with EDFs. Both the carbon and hydrogens present a local minimum close to the binding energy minimum.

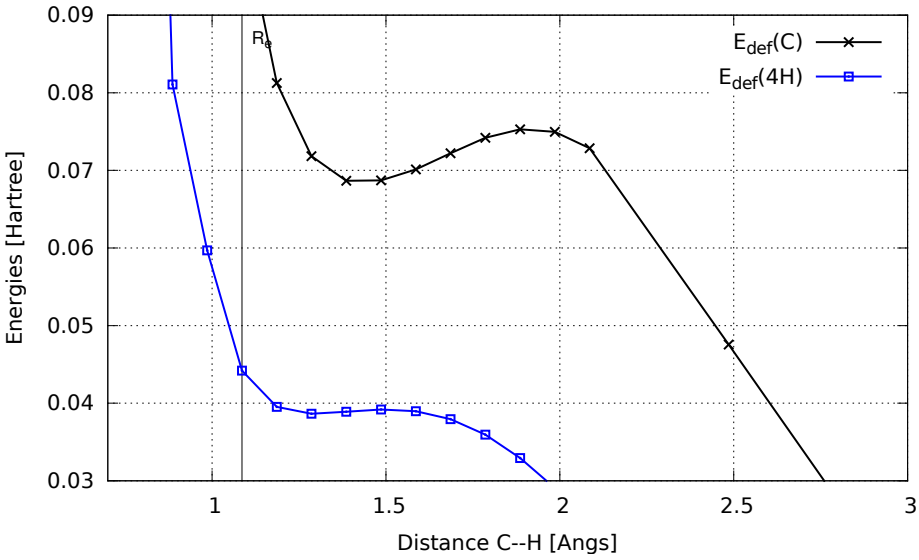


Figure 9: Carbon deformation energies with respect to the ^3P state, similarly the deformation energy of hydrogen takes as reference the ^2S fundamental state. The vertical line indicates the equilibrium geometry.

The approach of atoms has the effect of increasing delocalization between the two groups, carbon on one side and hydrogens on the other. If we consider the spinless EDF probabilities of finding a given number of electrons in the C and

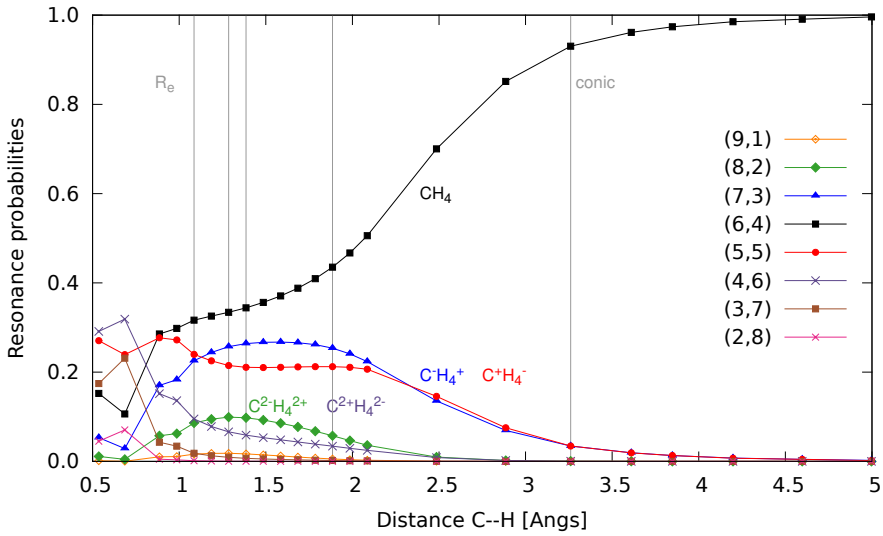


Figure 10: Spinless resonance structures of methane. Carbon constitutes the first group and equivalent hydrogens the other group, (n_C,n_{4H}) meaning the number of electrons of carbon and of the hydrogens respectively. Vertical lines indicate equilibrium geometry, minimum of the hydrogens deformation energy, minimum of the carbon deformation energy, hillock of the carbon deformation energy, and the conic intersection geometry respectively.

$_4$H subsystems, The "aufbau" structure (6,4) looses weight (figure 10) letting others, (7,3)[$C^-H_4^+$] and (5,5)[$C^+H_4^-$], increase. The purely covalent nature is broken at shorter distances, with an unbalanced weight of the newly populated structures (charge transfer) due to a difference in electronegativity of the atoms involved. Similarly, other resonant structures with even more delocalization are populated. Coincidentally, at the equilibrium geometry the polarization almost vanishes, in turn hydrogens become more attractive for electron delocalization. This condition allows us to avoid charge reorganization effects in the choice of the prepared state.

The (6,4) structure identifies the distribution of carbon that is isoelectronic with the isolated atom. If we would like now to describe it in an atomic-like fashion, it would be a mixture of the many atomic states. Restricting ourselves to intra-configuration excitations, a combination of the well known ^{3}P,^{1} D and ^{1}S carbon multiplets. In the dissociation limit the fundamental state of carbon is a ^{3}P with three degenerated M_S components. One component with an equal number of alpha and beta electrons $M_S = 0$ and two with one extra alpha $M_S = 1$ or beta $M_S = -1$ electron, respectively. The corresponding spin resonant structures are (3,3,2,2)[$M_S = 0$], (4,2,1,3)[$M_S = 1$] and (2,4,3,1)[$M_S = 1$], where the first two numbers describe the number of alpha and beta electrons in the carbon domain, and the second two the number of alpha and beta carbons in the remaining four hydrogen atoms. The weight of each component is obviously $1/3$. Our results are in tune with this fact (figure 11). As the distance is shortened the degeneracies are lost leading to the more important dynamic correlation. Even more, a higher multiplicity state is created at intermediate distances between bond dissociation and equilibrium that satisfies our expectations. The proposed bond formation mechanism relied on the formation of the ^{5}S state that is shown here as a (5,1,0,4) or (1,5,4,0) resonant structure. A non-intra-configurational state has been populated indicating that the aforementioned promotion has taken place. The existence of the ^{5}S state is localized in a specific range of geometries previous to bond formation.

It would be misleading to judge the weight of spin states taking spin resonance structures as reference because they contain various spin states and multiplicity has to be accounted. After this nuances are corrected we arrive at a much higher contribution by quintuplet states (figure 12). There is no doubt that the emergence of the quintuplet represents an essential step in the formation of the bonds, notwithstanding, it is not as predominant as it is usually assumed.

This hump has also been observed in electron correlation profiles. If two fragments do not interact as we bring them closer they will excite to their corresponding higher energy states. In the opposite situation, if they interact, the probability of the excited fragment states will increase smoothly (sigmoid). In the case of the methane formation, carbon opposes to interact a little bit and then all goes as expected.

The triplet transits from being absolutely dominant to the corresponding contribution of independent bonds without getting any fright. Singlets are mixed

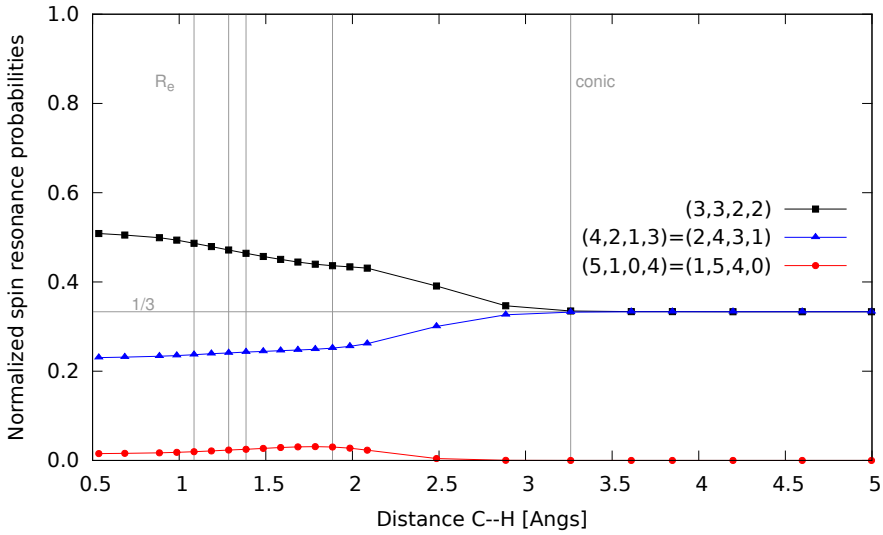


Figure 11: Spin resonance structures of methane that correspond to a neutral carbon, with notation $(n^{\alpha}_{C},n^{\beta}_{C},n^{\alpha}_{4H},n^{\beta}_{4H})$.

together in an unknown proportion. Singlets and quintuplet have the same proportion near the carbon deformation energy hillock. Whereas the quintuplet has its maximum peak there, the singlets continue increasing steadily.

The bonding energy $4 \cdot BE(CH) = 4 \cdot -92.8$ kcal/mol hinders a large relaxation of the fragments $E_R(CH_4) = 170$ kcal/mol. Remarkably the relaxation of hydrogens is not negligible: $E_R(H) = 27 kcal/mol$ in contrast with $E_R(C) = 62 kcal/mol$. In summary the intrinsic bond energy is $E_{int}(CH) = -128$ kcal/-mol. Our results are not far from a similar calculation of Cremer et al. that estimates a relaxation of 162 kcal/mol and IBE(C-H) of 140 kcal/mol. Notwithstanding, the source of relaxation is completely different. Carbon atoms do not promote completely and hydrogens relax more than expected.

At large distances we hope to find only distributions with single electrons in each orbital. Among those the Slater determinants with $M_S = \pm 2$ are not possible. So we are tied to 14 possible determinants, 6 with $M_S = 0$ couple with the determinants that generate the (3,3) resonant structure (probability 1/3). And sets of 4 determinants with $M_S = \pm 1$ conform the (1,3) and (3,1) resonances, each with probabilities 1/3. This means that not all determinants have the same probability: those with $M_S = 0$ have probability 1/18 and those of $M_S = \pm 1$ have 1/24.

4.4.2 Dinitrogen molecule N_2

The nitrogen atoms having all electrons unpaired in their ground states do not need to excite to a higher energy configuration to form three bonds, The

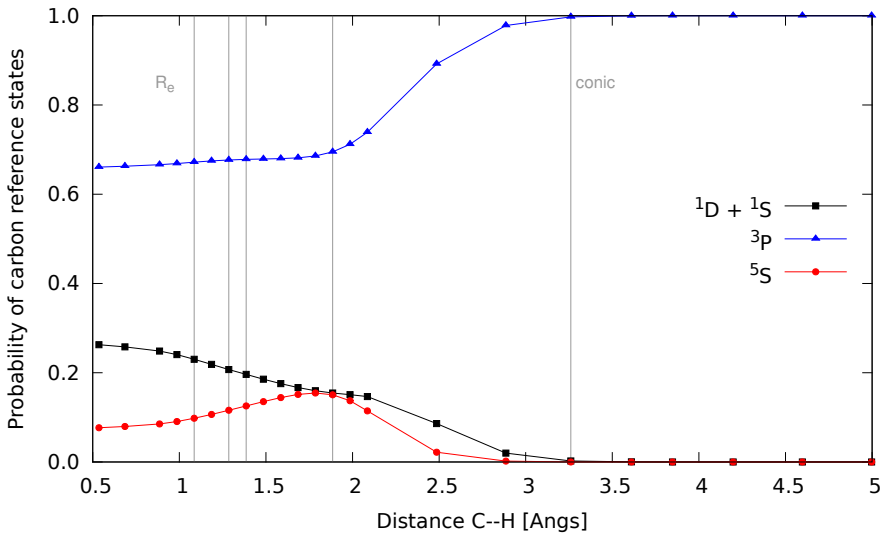


Figure 12: Probabilities of carbon reference states along the bonding process.

Table 3: A model of eight electrons and four independent bonds has probabilities that follow a B(8,1/2) binomial distribution. Probabilities are expressed as $p(6+\delta, 4-\delta)$ or $p(6-\delta, 4+\delta)$. The probability values for the condensed geometry ($d_{CH} = 0.8860$ angstrom) compare reasonably well.

	$\delta = 0$	$\delta = 1$	$\delta = 2$	$\delta = 3$
$p_{\text{binom.}}$	0.2734	0.2188	0.1094	0.0312
$p_{(\delta,-\delta)}$	0.2853	0.1700	0.1516	0.0101
$p_{(-\delta,\delta)}$	0.2853	0.2769	0.0574	0.0429

Table 4: Electron distribution functions including spin in the limit case of independent particles. The core of carbon is fixed ($1s^2$). Eight electrons are free to create four bonds, of those four belong to a group basin in the neutral resonant structure, half α and half β in a spin un-polarized structure. Thus the probabilities of the carbon structures are $p(2+\sigma, 2-\sigma) = p(2-\sigma, 2+\sigma)$. Calculated values are at $d_{CH} = 0.5361$ angstrom.

	$\sigma = 0$	$\sigma = 1$	$\sigma = 2$
$\mathfrak{h}(2-\sigma; 8, 4, 4)$	0.5143	0.2286	0.0143
$p_{(\sigma,-\sigma)}$	0.5085	0.2304	0.0153

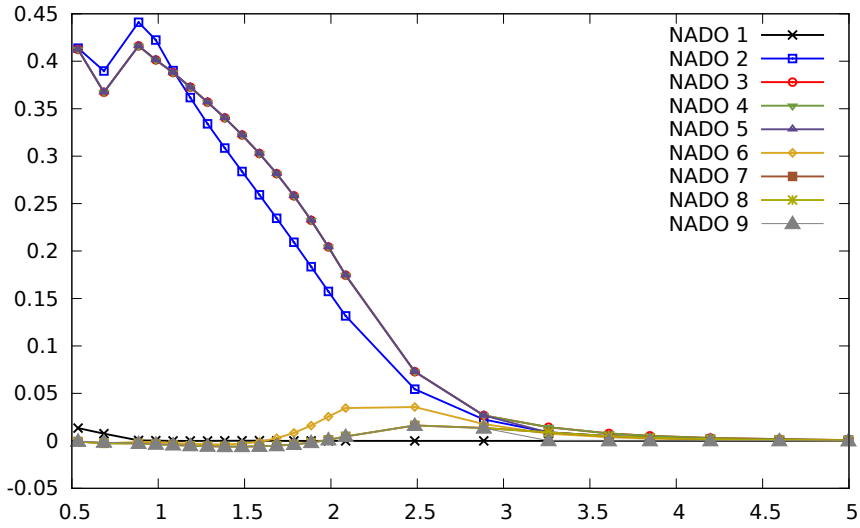


Figure 13: Natural adaptive occupation numbers of CH_4.

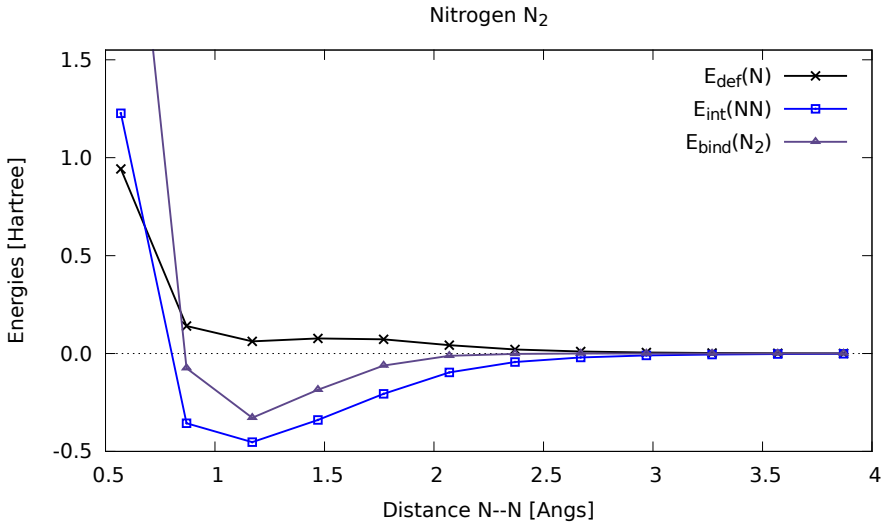


Figure 14: Nitrogen deformation energy with respect to the ^4S state, $E_{def}(N)$, interaction energy, $E_{int}(N,N)$, and binding energy of the molecule $E_{bind}(N_2)$.

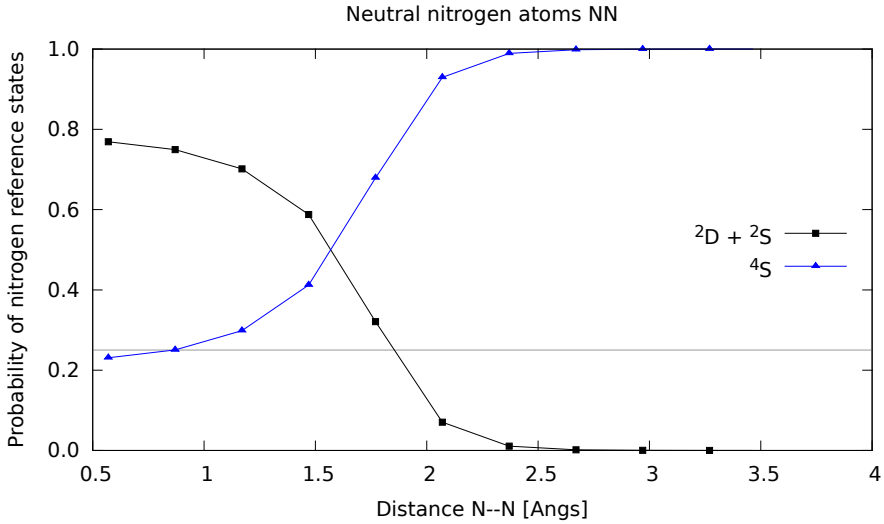


Figure 15: The electronic states of the fragment constituted by a nitrogen atom.

terms provided by the configuration $2s^2 2p^3$ suffice to create the bonding between nitrogen atoms. Therefore, there is no need to prepare the electronic configuration with an excitation to a higher energy state from a classical perspective. However, based on our previous results for methane, an increase of the weight of other intra-configurational states ^2S and ^2D is possible, but not of an inter-configurational state.

Table 5: A model of six electrons and three independent bonds has probabilities that follow a B(6,1/2) binomial distribution. Probabilities are expressed as $p(7+\delta, 7-\delta)$ or $p(7-\delta, 7+\delta)$ for this homonuclear case. Calculated values are at $d_{CH} = 0.5700$ angstrom.

	$\delta = 0$	$\delta = 1$	$\delta = 2$	$\delta = 3$
$p_{\text{binom.}}$	0.3125	0.2344	0.0938	0.0156
$p_{(\delta,-\delta)}$	0.3366	0.2386	0.0814	0.0112

Table 6: Electron distribution functions including spin in the limit case of independent particles. The core of each nitrogen is fixed ($1s^2$). Ten electrons are free to create three bonds, five of those belong to each basin in the neutral resonant structure, half α and half β in a spin un-polarized structure. Thus the probabilities of the nitrogen structures are $p(3+\sigma, 2-\sigma) = p(3-\sigma, 2+\sigma) = p(2+\sigma, 3-\sigma) = p(2-\sigma, 3+\sigma)$. Calculated values are at $d_{CH} = 1.4700$ angstrom.

	$\sigma = 0$	$\sigma = 1$	$\sigma = 2$
p_{hgeom}	0.3968	0.0992	0.0040
$p_{(\sigma,-\sigma)}$	0.3969	0.1031	0.0000

Spin quenching

As previously indicated for methane, dinitrogen has also a deformation energy minimum in the vicinity of the equilibrium structure (see figure 14). This behaviour was already discovered ([67], chapter 13).

One surprising feature of the equilibrium is that the doublets have larger contribution than the quadruplet, which is assumed to be the most important state. Those are statements which, although chemically appealing, are nevertheless misleading. The results follow almost directly from the independent particle model (see table 6).

N_2 is a prototypical molecule with a triple bond. One strong σ bond and two weaker π bonds. A point to be discussed is if it is observed that the creation of σ bonds takes place before than for π bonds, as it is indicated in the natural orbital occupations. With two-NAdOs we have a clear picture of bonding formation aside of the abstract molecular orbital reasoning commonly used for these explanations. Our results ratify the two step, σ and π, bonding formation scenario at the correlated level. A characteristic point to mark the half-bond creation is the inflection point. The inflection point for sigma bonding is 1.9338054 angstroms whereas it is 1.70300731 angstroms for π bonding formation. The contribution of sigma orbitals starts at long distances, before the π bonding increases due to correlation. As soon as π bonding appears the combination of 2s orbitals and p_z hybridize into sp type orbitals and couple to form four orbitals.

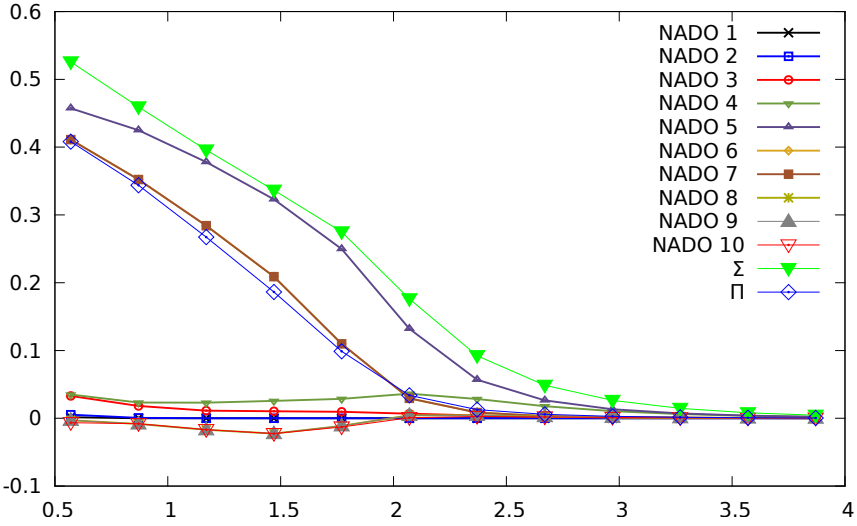


Figure 16: Natural adaptive occupation numbers of N_2.

4.4.3 Ethene/Ethylene

Methylene has two states very close in energy [5, 170, 171], a triplet $^3B_1$ and a singlet $^1A_1$. The former is only slightly more stable than the singlet.

The lowest energy state of ethene close to dissociation, i.e. 8 angstroms of separation between carbons, is a singlet with energy $-77.9445687087\,E_h$. This corresponds to the energy of two $^3CH_2$, $-38.9723018576\,E_h$ with a small size-consistency error. In the dissociation limit this state is degenerate with a quintuplet that is the result of the coupling of both triplets with maximum multiplicity. They do not intersect at least until 9.25 angstroms, the greatest separation we could achieve. Another singlet state was found, with higher energy, that correlates with the lowest energy CH_2 ($^1A_1$) + CH_2 ($^1A_1$) asymptote (of two $^1CH_2$ placed infinitely far apart). And it is size-consistent up to $1.76957550003465e-04\,E_h$.

The deformation energy of CH2 increases linearly as the bond is formed (see figure 17), as expected. The ground state of CH2 is a triplet so there should be no need to create and intermediate state in the course. This is confirmed by the probabilities of the electronic states (figure 18), although the relative weights of the states at equilibrium do not match with the standard assumption that the triplet will be predominant. Instead the singlet dominates with almost 60%. However, this is in line with our previous results wich indicated a sequence with hypergeometric probabilities (table 7).

The small energetic difference between the triplet and the singlet make for a small deformation energy of CH2 also, 31.6 kcal/mol, whereas the interaction energy is 211.2 kcal/mol. Then, most of the binding energy is involved in the sharing of electrons. The two main components are two NAdOs that can be interpreted as the sigma and pi bond (figure 19). Those are formed asynchronously, the sigma bond can start its creation at longer distances.

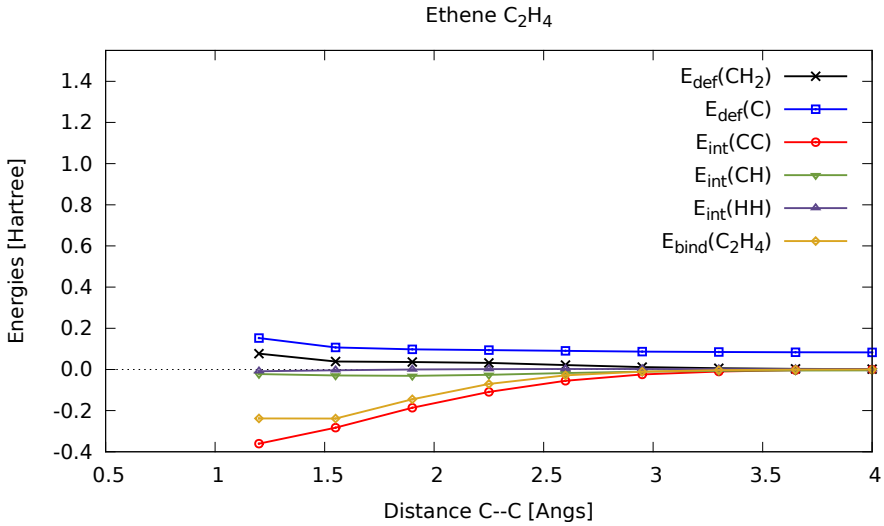


Figure 17: Energetic components of the C2H4 binding energy.

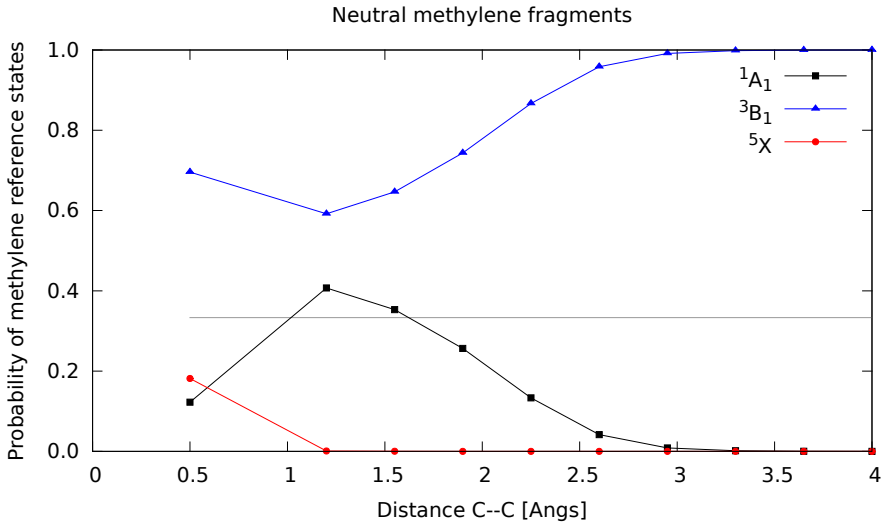


Figure 18: The spin states of the CH_2 fragment.

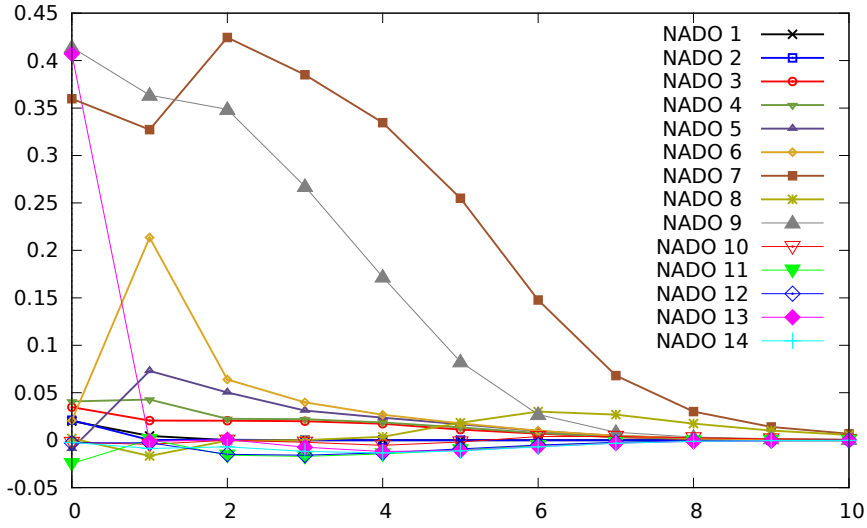


Figure 19: Contribution of CH_2 NAdOs to the formal double bond of ethylene. The graph depicts their natural occupation numbers. The nados 7 and 9, which correspond to the orbitals that most closely resemble the sigma and pi bonding orbitals have the greatest contribution to bonding. In this case, as we have seen for dinitrogen, the bonding is created at two states, differentiating sigma and pi components.

Table 7: Electron distribution functions including spin in the limit case of independent particles. The core of each carbon is fixed ($1s^2$). Twelve electrons are free to create two bonds, of those six belong to each basin in the neutral resonant structure, half α and half β in a spin un-polarized structure. Thus the probabilities of the nitrogen structures are $p(3+\sigma, 3-\sigma) = p(3-\sigma, 3+\sigma)$. Calculated values are at $d_{CC} = 0.5$ angstrom.

	$\sigma = 0$	$\sigma = 1$	$\sigma = 2$
p_{hgeom}	0.4329	0.2435	0.0390
$p_{(\sigma,-\sigma)}$	0.4150	0.2561	0.0363

4.4.4 Ethyne/Acetylene

Methylidyne (CH) has a minimum in the deformation energy close to the equilibrium geometry. So the internal state finds some stabilization that favors the equilibrium. At this point we can figure outh how the probabilities of the states will be. Being the ground state of acetylene a doublet (methylylidene) we can expect in advance that the quartet (methanetriyl) will be generated in some extent during the reaction, in contrast to dinitrogen where the quartet is already the ground state. However, at equilibrium the quartet should come back to a lower probability. In fact that is what is seen, albeit the weight that the quartet reachtes at half bonding formation is notably greater than what we would expect from inpecting the antecedent of methane. We also remark how abruptly this state is formed at around 2.5 angstroms of carbon-carbon distance.

If we take a look at the probabilities of both dinitrogen and acetylene reference states we can realize that close to equilibrium the weights of the doublets and quartets are roughly the same, the electronic combinations of the independent model dictate the behaviour.

4.5 CONCLUSIONS

In this chapter we have found that bonding energies in polyatomic molecules cannot be measured unless some approximations are made. Even for diatomic molecules a reference needs to be fixed. Assigning a reference we can obtain an indicator of bond strength, but there is no computational method to realize such measure. This is where IQA partitioning plays a fundamental role establishing a solid foundation for the concept of intrinsic bond energy. Furthermore, this quantity is available for any pair of atoms that we conceive, without being restricted to the shortcomings of experimental accessibility. Together with the rest of QCT tools we can characterize the reference, i.e. by identification of the electronic state that better represents the valence state of the fragments that unite to form the bond. It is deducted from the results that high spin

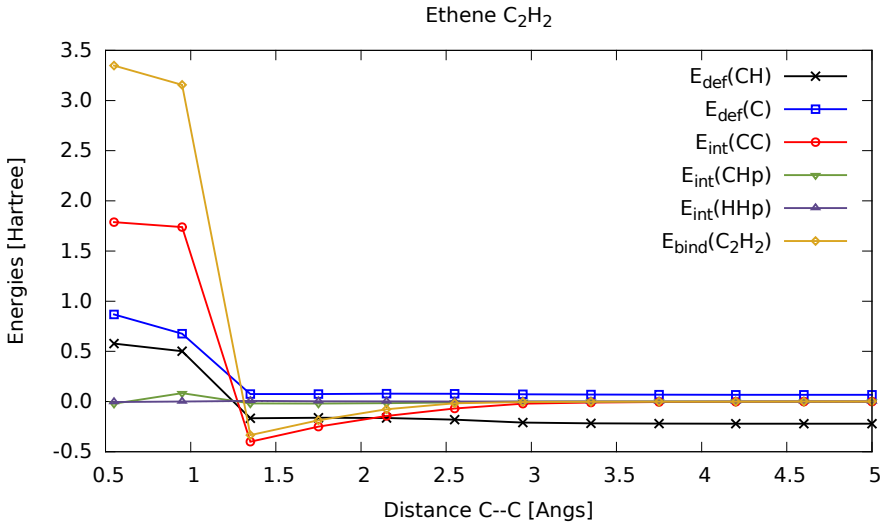


Figure 20: Energetic components of the C2H2 binding energy.

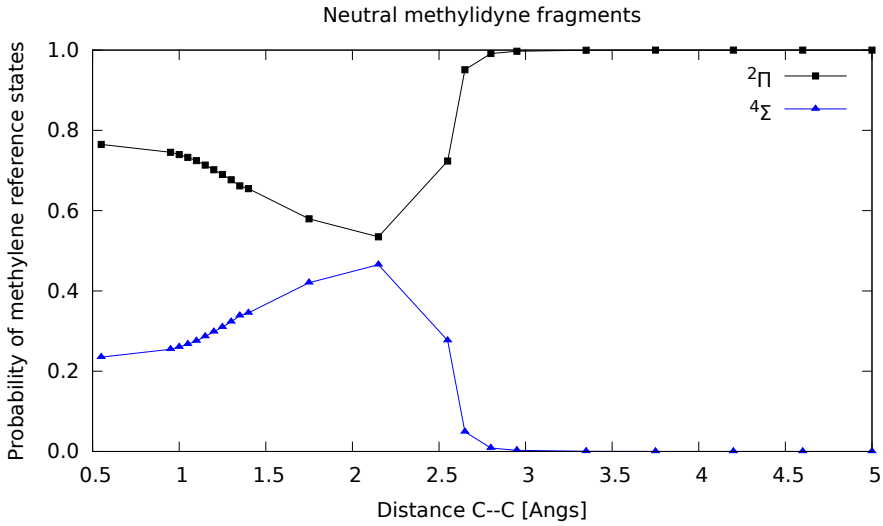
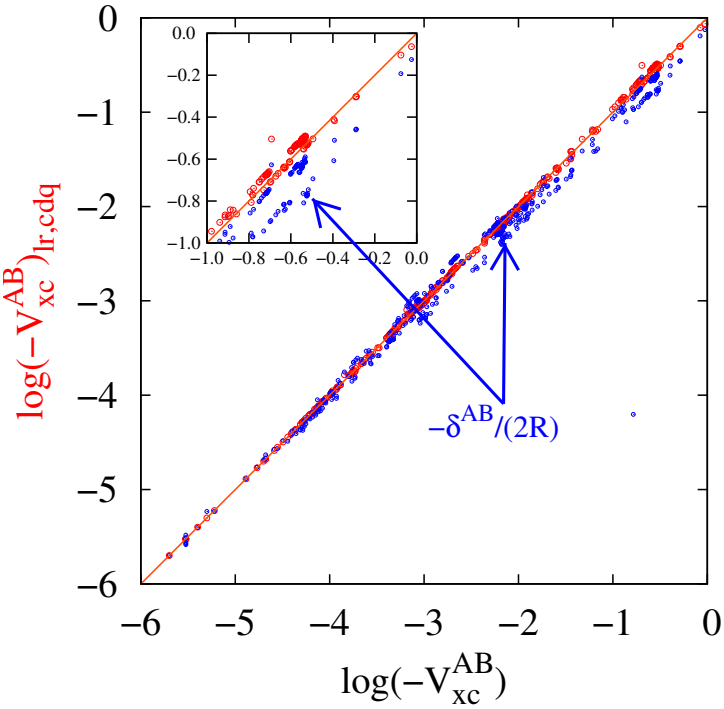


Figure 21: The spin states of the CH fragment.

electronic configurations are less prone to happen than it is often claimed. This indicates us that we should be particularly careful when attributing valence states without a clear justification. Apart from this there are no great changes, with respect to standard assumptions, in the electronic behaviour. Also, we have appreciated the overestimation of bonding energies because interactions between atoms that are not supposed to be bonded are by mistake included in the bond energy estimation.

5 MULTIPOLAR EXPANSION OF THE EXCHANGE-CORRELATION INTERACTION ENERGY

CONTENTS



Interatomic or interfragment covalent energies in real space, as measured by the interacting quantum atoms (IQA) exchange-correlation energies (V_{xc}^{AB}) are shown to be well-approximated by a multipolar approximation if terms up to the charge-quadrupole interaction are retained (cdq). The cdq approximation improves considerably the performance of the zeroth-order approximation, in which V_{xc} is equal to the bond order (delocalization index, δ^{AB}) over the interatomic distance.

Interatomic exchange-correlation energies correspond to the covalent energetic contributions to an interatomic interaction in real space theories of the chemical bond, but their widespread use is severely limited due to their computationally intensive character. In the same way as the multipolar (*mp*) expansion is customary used in biomolecular modelling to approximate the classical Coulomb interaction between two charge densities $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$, we examine in this work the *mp* approach to approximate the interatomic exchange-correlation (*xc*) energies of the Interacting Quantum Atoms method. We show that the full *xc mp* series is quickly divergent for directly bonded atoms (1–2 pairs) albeit it works reasonably well most times for 1–n ($n > 2$) interactions. As with conventional perturbation theory, we show numerically that the *xc* series is asymptotically convergent and that, a truncated *xc mp* approximation retaining terms up to $l_1 + l_2 = 2$ usually gives relatively accurate results, sometimes even for directly bonded atoms. Our findings are supported by extensive numerical analyses on a variety of systems that range from several standard hydrogen bonded dimers to typically covalent or aromatic molecules. The exact algebraic relationship between the monopole-monopole *xc mp* term and the inter-atomic bond order, as measured by the delocalization index of the Quantum Theory of Atoms in Molecules, is also established.

5.1 INTRODUCTION

The role of the quantum mechanical exchange-correlation (*xc*) energy as the basic glue binding together atoms and molecules has been clearly stressed in the past [96]. In the chemical literature, however, this insight is less well known. Although exchange-correlation functionals, for instance, are the essential ingredients in modern implementations of Density Functional Theory (DFT) [89], not much work has been devoted to examine the importance of the exchange-correlation energy itself in the theory of chemical bonding from the DFT viewpoint [14].

Actually, almost all that is known about the chemical relevance of the *xc* energy has been derived in the last decade through the study of bonding in real or position space [51]. With this term we gather together a number of techniques that are being actively explored [32, 90, 143] which use orbital invariant reduced densities (or density matrices) to develop a new paradigm that may one day replace the standard molecular orbital approach [69]. Usually, these techniques use a partition of real space into regions endowed with chemical meaning, be them atoms, bonds, cores, lone pairs, etc. In many cases, the space is divided using the topology induced by the gradient field of an orbital invariant scalar, like the electron density (which gives rise to the atomic partioning of the quantum theory of atoms in molecules (QTAIM) developed by Bader and coworkers [10], or the electron localization function (that isolates core, bond

and lone pair regions) [18, 174]. When this topological tools are used we say that we are under the Quantum Chemical Topology umbrella [154].

In the context of the QTAIM/QCT, we proposed a number of years ago an exact, general decomposition of the total molecular energy E into atomic and interatomic terms that we called the interacting quantum atoms (IQA) approach [51]. All the expectation values of the standard Coulomb Hamiltonian that make up E are written in IQA as a sum of domain contributions, and E is obtained by adding atomic self-energies, which tend to the free atomic energies when the atoms that interact are sufficiently far apart, and pairwise additive interaction energies. The latter are composed of a classical term that depends only on classical electrostatic contributions, and an exchange-correlation energy, V_{xc} which accounts for the quantum mechanical effects. As we and others have shown over the years [73, 109], the classical part of the interaction measures its ionic component, while the V_{xc} energy is to be associated with its covalent counterpart.

In these years, the interatomic *xc* energy has become an important ingredient of any quantitative account of chemical bonding in position space [33, 67]. For instance, it has been shown to be intimately related to the appearance of the bond critical points of the QTAIM, leading to the concept of priviledged exchange-correlation channels [148]. It has also been used to reconstruct molecular graphs from purely energetic quantities [64], to shed light on new concepts like halogen bonding [15, 16], to recover stereolectronic effects [144], or to find new long-range electronic anomalies [63].

Interatomic V_{xc} energies are intimately linked to the delocalization or shared electron delocalization indices (DIs) used in the QTAIM, defined almost 40 years ago by Bader and Stephens [9]. These are obtained by directly integrating the *xc* density of very two different atomic domains and measure the number of shared pairs of electrons between them. They have been successfully used as real space generalization of the bond order concept, reducing to the Wiberg-Mayer [118, 190] bond orders if atomic domains are imagined to collapse onto their nuclei. In a sense, V_{xc}'s are the energetic counterparts of DIs, and both have been empirically found to correlate very well when a given couple of atoms is examined in different molecular environments.

The computational complexity of obtaining DIs is considerably smaller than that of calculating V_{xc}'s, since the former may be factorized into sums of products of atomic overlap matrices (3D numerical integrals), while the latter need, in principle, very costly 6D quadratures. Thus, if we are not interested in very accurate results, but only in semi-quantitative estimations of covalent energies, any procedure that might approximate the V_{xc} values in terms of cheaper to compute quantities like the DIs should be welcome. That procedure was initially examined by Rafat and Popelier [157], that wrote each interatomic Hartree-Fock (HF) V_{xc} interaction as a multipolar series, exploring the convergence of this series in different closed-shell molecules computed at this level of theory. In this work, we generalize their algebraic formalism to multideterminant wavefunctions. This generalization is possible thanks to the use of

the monadic diagonalization of the exchange-correlation density [146], customary used within the IQA methodology [51]. Our expressions converge to those of Ref. 157 when HF exchange-correlation densities are employed in the calculation. We will show that, regardless the type of calculation, the monopole-monopole term of the multipolar *xc* interaction between two atoms A and B of the molecule coincides with that of Rafat and Popelier, being equal to $-\delta^{AB}/(2R)$ (R is the AB internuclear distance), and that the series is usually divergent although many times asymptotically convergent. Moreover, our results clearly establish in what conditions V_{xc} can be safely approximated by a truncated series, and how in some situations retaining up to the charge-quadrupole terms may give reasonable results even for directly bonded atoms. In the latter cases, the use of the crudest approach $V_{xc}^{AB} \sim -\delta^{AB}/(2R)$ is even preferable to using the multipolar expansion up to very high order.

We will first consider the multipolar expansion of V_{xc}, including a short account of the IQA methodology. Then we will turn to examine how the series converges or diverges for a number of selected systems.

5.2 MULTIPOLAR EXPANSION OF V_{xc}^{AB}

In this section, we briefly describe the Interacting Quantum Atoms (IQA) method and the role played by the exchange-correlation (*xc*) interaction in this energy partition method (Subsection 5.2.1), the exact computation of this interaction (Subsection 5.2.2), and its multipolar approximation with or without truncating the expansion of the angular momentum series (Subsection 5.2.3). It is worth noting that the experience gained to date with the IQA method, both by us and by other groups, clearly indicates that the magnitude of V_{xc}^{AB} correlates very well with the degree of covalency between the pair of atoms A and B as measured by means of the delocalization index defined by Bader and Stephens, and weighted through the inverse of the distance between both atoms. As we will see, this correlation would be perfect as long as the crudest multipolar approximation to V_{xc}^{AB} (consisting in truncating the multipolar series in the term $l_1 = m_1 = l_2 = m_2 = 0$) were exact.

5.2.1 The Interacting Quantum Atoms (IQA) method

The IQA method [51] is a real space energetic partition inspired in the Quantum Theory of Atoms in Molecules (QTAIM) that focuses on domain-averaged integrated quantities. The total energy in this approach is given by

$$E = \sum_A T_A + V_{en}^{AA} + V_{ee}^{AA} + \sum_{A>B} V_{nn}^{AB} + V_{en}^{AB} + V_{en}^{BA} + V_{ee}^{AB} \quad (5\text{–}1)$$

$$= \sum_A E_{self}^A + \sum_{A>B} E_{int}^{AB}. \quad (5\text{–}2)$$

where A runs over all the atoms in the molecule, $V_{nn}^{AB} = Z^A Z^B / R_{AB}$ is the repulsion between the nuclei A and B, $V_{en}^{AB} = -Z^B \int_{\Omega_A} d\mathbf{r}_1 \ \rho(\mathbf{r}_1) r_{1B}^{-1}$ is the nuclear attraction of the electrons within the basin of A (Ω_A) to the nucleus B, and V_{ee}^{AB} is the total electron repulsion between Ω_A and Ω_B. The latter is given by $V_{ee}^{AB} = J^{AB} + V_{xc}^{AB}$ where

$$J^{AB} = \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \, r_{12}^{-1} \, \rho(\mathbf{r}_1) \, \rho(\mathbf{r}_2), \tag{5–3}$$

is the classical or Coulomb electron-electrons repulsion, and

$$V_{xc}^{AB} = \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2 \, r_{12}^{-1} \, \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2), \tag{5–4}$$

where $\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$ is the exchange-correlation (*xc*) density, is the purely quantum-mechanical electron-electron *xc* interaction, which is the main subject of this work. In this way,

$$E_{int}^{AB} = \left(V_{nn}^{AB} + V_{en}^{AB} + V_{en}^{BA} + J^{AB} \right) + V_{xc}^{AB} = V_{cl}^{AB} + V_{xc}^{AB}. \tag{5–5}$$

The term E_{self}^{A} in (5–2) collects all the energetic components affecting exclusively to the atom A while E_{int}^{AB} represents the full interaction energy between atoms A and B, that is made of the full electrostatic or classical interaction (V_{cl}^{AB}) and the quantum-mechanical part (V_{xc}^{AB}). The expression (5–2) is valid, not only for the IQA methodology, but also for other energetic partitions, such as a recently proposed one inspired in the IQA method, although using a fuzzy partition of the space and localized molecular orbitals (MO) [49]. Mayer and Hamza have also dealt with the exchange component in equation (5–4) in the framework of a Hilbert space partition instead of the real space QTAIM partition we use here [122].

5.2.2 The exact *xc* interaction energy

Over the years, it has become clear that the magnitude of V_{xc}^{AB} measures the degree of covalency of the chemical bond between the atoms A and B. The more negative its value, the bigger the bond order between the two atoms and vice versa [109, 143, 148]. Their values have been recently proposed as a novel solution to the problem of assigning a molecular graph to a collection of nuclei [63](i.e. how to draw a molecular structure). In the IQA approach, this term is exactly computed as follows. First, we use the fact that for both single- (1-det) and multi-determinant wavefunctions built in with real MOs ϕ_i, $\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$ can be written as

$$\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i,j,k,l}^{M} \lambda_{ijkl} \, \phi_i(\mathbf{r}_1) \, \phi_j(\mathbf{r}_1) \, \phi_k(\mathbf{r}_2) \, \phi_l(\mathbf{r}_2), \tag{5–6}$$

where M is the number of partially or fully occupied MOs, and λ_{ijkl} is a symmetric matrix in the (i,j) and (k,l) pairs. Defining a set of coefficients,

$$\epsilon_{ijkl} = \lambda_{ijkl} + \lambda_{jikl}(1-\delta_{ij}) + \lambda_{ijlk}(1-\delta_{kl}) + \lambda_{jilk}(1-\delta_{ij})(1-\delta_{kl}), \quad (5\text{–}7)$$

where δ_{ij} is the Kronecker symbol ($\delta_{ij} = 1$ for $i = j$, $\delta_{ij} = 0$ for $i \neq j$) we may write an (i,j),(k,l) symmetric simpler expression,

$$\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2) = \sum_{i\geqslant j,k\geqslant l}^{M} \epsilon_{ijkl}\phi_i(\mathbf{r}_1)\,\phi_j(\mathbf{r}_1)\,\phi_k(\mathbf{r}_2)\,\phi_l(\mathbf{r}_2). \quad (5\text{–}8)$$

Using the basis of products of MOs, $\{\phi_i(\mathbf{r})\phi_j(\mathbf{r}), i \geqslant j\}$, that contains $M(M+1)/2$ members, we diagonalize equation (5–8), and get: [146]

$$\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2) = \sum_{i\geqslant j}^{M} \eta_{ij} f_{ij}(\mathbf{r}_1) f_{ij}(\mathbf{r}_2), \quad (5\text{–}9)$$

where the f_{ij} eigenfunctions are linear combinations of the above products. The ϵ matrix may be easily computed from the explicit form of a given calculated wavefunction. For closed-shell 1-det wavefunctions (and formally also for a Kohn-Sham determinant) $M = N/2$, where N is the number of electrons, the ϵ matrix is already diagonal in the (i,j) and (k,l) pairs, each eigenvector is the product of two MOs, $f_{ij} = \phi_i\phi_j$, and the η_{ij} eigenvalues are simply $\eta_{ii} = -2$ and $\eta_{ij} = -4$ $(i \neq j)$. Using (5–9) in the expression of V_{xc}^{AB} one gets

$$V_{xc}^{AB} = \sum_{i\geqslant j}^{M} \eta_{ij} K_{ij}^{AB}, \quad \text{where} \quad (5\text{–}10)$$

$$K_{ij}^{AB} = \int_{\Omega_A} d\mathbf{r}_1 \int_{\Omega_B} d\mathbf{r}_2\; r_{12}^{-1}\; f_{ij}(\mathbf{r}_1) f_{ij}(\mathbf{r}_2). \quad (5\text{–}11)$$

The integrals (5–3) and (5–11) can be computed numerically and (in principle) exactly, i.e without invoking any approximation such as the multipolar expansion, by means of the bipolar expansion as described in Ref. 142. Notice that using the Fock-Dirac exchange from Kohn-Sham determinants is an approximation that has no rigorous justification.

5.2.3 The multipolar approach for V_{xc}^{AB}

Comparing equation (5–11) with equation (5–3) for the Coulomb repulsion it is evident that if J^{AB} is approximated making use of physically reasonable arguments that are also valid for K_{ij}^{AB}, the steps to approximate the latter will be the same used for J^{AB}. The long-range or multipolar approximation (MP) to J^{AB}, given by

$$J_{lr}^{AB} = \sum_{l_1 m_1}^{\infty} \sum_{l_2 m_2}^{\infty} C_{l_1 m_1 l_2 m_2}(\hat{R}) \frac{Q_{l_1 m_1}^{\Omega_A} Q_{l_2 m_2}^{\Omega_B}}{R^{l_1+l_2+1}}, \quad (5\text{–}12)$$

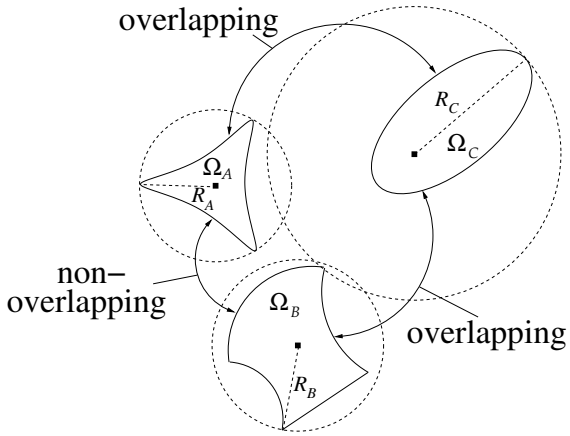


Figure 22: Schematic representation of overlapping and non-overlapping regions.

where m_1 (m_2) runs from $-l_1$ ($-l_2$) to $+l_1$ ($+l_2$), $\mathbf{R} = (\mathbf{R}_B - \mathbf{R}_A) \equiv (R, \hat{R})$, with $R = |\mathbf{R}_B - \mathbf{R}_A|$ and $\hat{R} \equiv (\theta_B - \theta_A, \phi_B - \phi_A)$ (see figure 56) is the position vector of the B center with respect to the A center, $C_{l_1 m_1 l_2 m_2}(\hat{R})$ are known coefficients, Q_{lm}^{Ω} are the spherical atomic multipoles, defined as

$$Q_{lm}^{\Omega} = N_l \int_{\Omega} r^l \, S_{lm}(\hat{r}) \, \rho(\mathbf{r}) \, d\mathbf{r}, \tag{5–13}$$

$N_l = \sqrt{4\pi/(2l+1)}$, and $S_{lm}(\hat{r})$ are real spherical harmonics (see Appendix) is exact when the basins Ω_A and Ω_B are non-overlapping (See figure 22 and the definition of overlapping and non-overlapping regions below). Equation (5–12) is the same used by Popelier *et. al.* [92, 155, 156] in their discussion of the multipolar expansion for the diatomic Coulomb repulsion. Retaining only terms with $l_1 \leqslant 1$ and $l_2 \leqslant 1$ in this equation one has

$$J_{lr,cd}^{AB} \simeq \frac{Q^A Q^B}{R} - Q^A \frac{\vec{\mu}^B \cdot \mathbf{R}}{R^3} + Q^B \frac{\vec{\mu}^A \cdot \mathbf{R}}{R^3} + \frac{1}{R^3}\left(\vec{\mu}^A \cdot \vec{\mu}^B - 3\frac{(\vec{\mu}^A \cdot \mathbf{R})\,(\mathbf{R} \cdot \vec{\mu}^B)}{R^2}\right), \tag{5–14}$$

where $Q^{\Omega} = \int_{\Omega} \rho(\mathbf{r}) d\mathbf{r}$ and $\vec{\mu}^{\Omega} = \int_{\Omega} \mathbf{r}\rho(\mathbf{r}) d\mathbf{r}$ are the total electron charge and the dipole moment of the Ω region, respectively. The first, second plus third, and fourth terms of (5–14) correspond to the charge-charge (cc), charge-dipole (cd), and dipole-dipole (dd) interactions, respectively. We should note that the second and third terms have opposite signs.

If the same approximation is used for K_{ij}^{AB}, $\left(V_{xc}^{AB}\right)_{lr}$ becomes

$$\left(V_{xc}^{AB}\right)_{lr} = \sum_{l_1 m_1}^{\infty} \sum_{l_2 m_2}^{\infty} C_{l_1 m_1 l_2 m_2}(\hat{R}) \frac{\delta^{AB}_{l_1 m_1, l_2 m_2}}{R^{l_1+l_2+1}}, \quad \text{where} \tag{5–15}$$

$$\delta^{AB}_{l_1 m_1, l_2 m_2} = \sum_{i \geqslant j}^{M} \eta_{ij} q^{\Omega_A}_{ij, l_1 m_1} q^{\Omega_B}_{ij, l_2 m_2}, \quad \text{and} \tag{5–16}$$

$$q^{\Omega}_{ij, lm} = N_l \int_{\Omega} r^l \, S_{lm}(\hat{r}) f_{ij}(\mathbf{r}) d\mathbf{r}. \tag{5–17}$$

It is important to stress that, similarly to $(J^{AB})_{lr}$, the expression (5–15) for $(V_{xc}^{AB})_{lr}$ provides the exact xc interaction when the atomic basins Ω_A and Ω_B do not overlap (figure 22). In the present context these two basins are non-overlapping because the two spheres of radii R_A and R_B, centered at the origin of Ω_A and Ω_B, respectively, do not intersect each other, being R_A (R_B) the maximum distance from the origin of the basin to the surface of Ω_A (Ω_B). On the contrary, Ω_A and Ω_C are overlapping despite that none point inside Ω_A belongs also to Ω_C and viceversa. When the non-overlapping condition is not met, the current expressions for $(J^{AB})_{lr}$ and $(V_{xc}^{AB})_{lr}$ are only conditionally convergent. We will see different examples of this in Section 5.4.

The function $N_l r^l S_{lm}(\hat{r})$ is 1 for $l = m = 0$, (y, z, x) for $l = 1$ and $m = (-1, 0, +1)$, and $(\sqrt{3}xy, \sqrt{3}yz, \frac{1}{2}(3z^2 - r^2), \sqrt{3}xz, \frac{\sqrt{3}}{2}(x^2 - y^2))$ for $l = 2$ and $m = (-2, -1, 0, +1, +2)$. If, as in the case of J^{AB}, only terms with $l_1 \leqslant 1$ and $l_2 \leqslant 1$ are included, $(V_{xc}^{AB})_{lr}$ becomes

$$\begin{aligned}\left(V_{xc}^{AB}\right)_{lr,cd} \simeq \sum_{i,j} \eta_{ij} \Bigg[& \frac{q^{\Omega_A}_{ij} q^{\Omega_B}_{ij}}{R} - q^{\Omega_A}_{ij} \frac{\vec{\mu}^{\Omega_B}_{ij} \cdot \mathbf{R}}{R^3} + q^{\Omega_B}_{ij} \frac{\vec{\mu}^{\Omega_A}_{ij} \cdot \mathbf{R}}{R^3} \\ & + \frac{1}{R^3} \left(\vec{\mu}^{\Omega_A}_{ij} \cdot \vec{\mu}^{\Omega_B}_{ij} - 3 \frac{(\vec{\mu}^{\Omega_A}_{ij} \cdot \mathbf{R}) \, (\mathbf{R} \cdot \vec{\mu}^{\Omega_B}_{ij})}{R^2} \right) \Bigg],\end{aligned} \tag{5–18}$$

where $q^{\Omega}_{ij} \equiv q^{\Omega}_{ij,00} = \int_{\Omega} f_{ij}(\mathbf{r}) d\mathbf{r}$, and

$$\vec{\mu}^{\Omega}_{ij} \equiv (q^{\Omega}_{ij,1-1}, q^{\Omega}_{ij,10}, q^{\Omega}_{ij,1+1}) = \int_{\Omega} \mathbf{r} f_{ij}(\mathbf{r}) d\mathbf{r}. \tag{5–19}$$

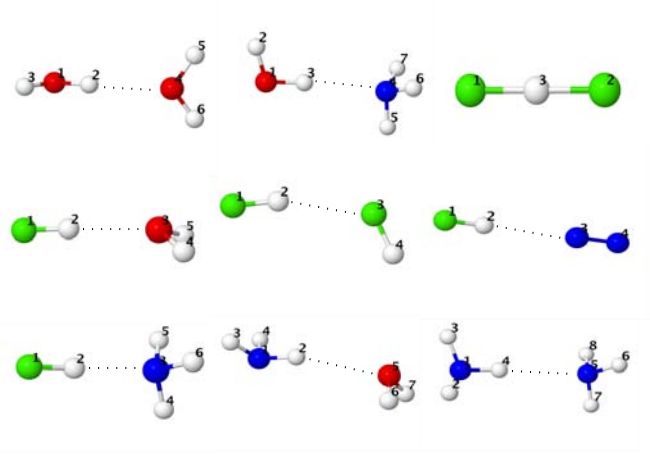
If terms with $(l_1 = 0, l_2 = 2)$ and $(l_1 = 2, l_2 = 0)$ are also included, the extra contribution

$$(V_{xc}^{AB})_{lr,cq} = \sum_{i,j}^{M} \frac{\eta_{ij}}{R^3} \sum_{m=-2}^{+2} q_{2m}(\hat{R}) \left[q^{A}_{ij} q^{B}_{ij,2m} + q^{B}_{ij} q^{A}_{ij,2m} \right] \tag{5–20}$$

must be added to (5–18). The cq subscript in equation (5–20) stands for *charge-quadrupole* interactions. The improved expression for $(V_{xc}^{AB})_{lr}$ is then

$$(V_{xc}^{AB})_{lr,cdq} = (V_{xc}^{AB})_{lr,cd} + (V_{xc}^{AB})_{lr,cq}. \tag{5–21}$$

Figure 23: Hydrogen bond systems studied in this work. Hydrogen, nitrogen, oxygen, and fluorine atoms are represented as growing size spheres, respectively.



The physical meaning of q_{ij}^{Ω} and $\vec{\mu}_{ij}^{\Omega}$ are easy to grasp. If we consider the particular case of their diagonal expressions ($i = j$) for a 1-det wavefunction, $f_{ii}(\mathbf{r}) = \phi_i^2(\mathbf{r})$, so that q_{ii}^{Ω} is the electron charge of the orbital distribution $\phi_i^2(\mathbf{r})$ within the Ω region, and $\vec{\mu}_{ii}^{\Omega}$ the dipole moment of Ω due to this distribution. For this reason, q_{ij}^{Ω} and $\vec{\mu}_{ij}^{\Omega}$ may be called orbital overlap charge and orbital overlap dipole, respectively. At the Hartree-Fock (HF) level, the q_{ij}^{Ω}'s coincide with the Atomic Overlap Matrix (AOM) elements of the QTAIM, $q_{ij}^{\Omega} \equiv \langle\phi_i|\phi_j\rangle_{\Omega} = S_{ij}^{\Omega}$. However, given that $f_{ij}(\mathbf{r})$ at the correlated level is a linear combination of $\phi_i(\mathbf{r})\phi_j(\mathbf{r})$ products, q_{ij}^{Ω} in this case is a linear combination of AOM elements. Nevertheless, for both types of wavefunctions $-2\delta_{00,00}^{AB}$ coincides with δ^{AB}, the so-called delocalization index (DI) between the atoms A and B

$$-2\delta_{00,00}^{AB} = \delta^{AB} = -\sum_{i \geqslant j}^{M} 2\eta_{ij} S_{ij}^{\Omega_A} S_{ij}^{\Omega_B}, \tag{5–22}$$

so that the leading term of $(V_{xc}^{AB})_{lr}$ ($l_1 = l_2 = m_1 = m_2 = 0$) can be written as

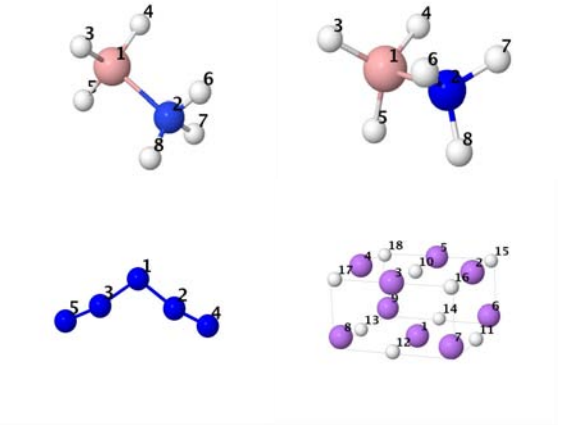
$$R^{-1} \sum_{i,j} \eta_{ij} q_{ij}^{\Omega_A} q_{ij}^{\Omega_B} = -\delta^{AB}/(2R). \tag{5–23}$$

The above equation is behind the good existing correlation between the values of V_{xc}^{AB} and δ^{AB} for a large collection of AB couples in many systems. The present derivation shows that the proportionality between V_{xc}^{AB} and δ^{AB} is modulated by the inverse of the distance between the nuclei of both atomic basins.

5.3 SYSTEMS AND COMPUTATIONAL DETAILS

All the calculations of this work have performed with our PROMOLDEN code [107]. This program allows the exact computation [142, 146] (i.e., without suffering the convergence problems inherent to the multipolar series expansion) of V_{xc}^{AB} as well as the full (lr) and truncated (lr,cd) and (lr,cdq) multipolar approximations described in Section 5.2. For brevity, only the exact, and the (lr) and (lr,cdq) numbers will be given in the tables. The errors plotted in the figures are defined as $[(V_{xc}^{AB})_{method} - (V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}| \times 100$, where method =(lr), (lr,cd), or (lr,cdq). The studied systems include several standard hydrogen bonded (HB) dimers (figure 23), the staggered BH_3NH_3, eclipsed BH_3NH_3, N_5^+, and Li_9H_9 molecules (figure 24), eleven molecules derived from saturated hydrocarbons by substituting C or H atoms by Be, B, N, O, F atoms, plus the benzene molecule (figure 25), the saturated hydrocarbons ethane, propane, butane, and pentane (figure 26), and the phenol dimer (figure 27). The labels of the atoms in the tables are those defined in these figures. For simplicity, the molecular orbitals required for evaluating the exchange-correlation density of equation (5–8) have been obtained through restricted Hartree-Fock (RHF) calculations at the corresponding equilibrium geometries with basis sets of quality 6-311G(d,p) or higher. However, since our results in this paper stem from the algebraic properties of the multipolar expansion, we do not expect significant changes neither in the numerical results nor in the subsequent discussion when using more accurate wavefunctions or the approximate data coming from Kohn-Sham determinants in the computation of the *xc* interactions. To prove the validity of this assertion, we will compare the V_x^{AB} energies obtained for staggered ethane in a CAS[14,14] calculation (Complete Active Space calculation with all except the carbon 1s electrons distributed into 14 valence orbitals) with the RHF results. All other *xc* interactions except those of the above CAS calculation lack a correlation energy component, being thus pure exchange contributions that should be more properly labelled V_x^{AB}. However, since all the expressions in Section 5.2 are valid for general wavefunctions the original name will be used hereinafter. The sums over l_1 and l_2 in equation (5–15) were truncated at $l_1^{max} = l_2^{max} = 8$, so that terms up to a range $L = l_1^{max} + l_2^{max} + 1 = 17$ were included in the multipolar expansion. Since QTAIM domains are usually finite and quite irregular, very fine radial and angular grids are needed to carry out the 6D numerical integrations. Here, we have systematically considered a $\beta-$sphere around each atom, with a radius equal to 60-90% the distance of its nucleus to the closest bond critical point, and employed high quality Lebedev angular and radial grids, with $(5810, 512)$ and $(194, 400)$ points outside and inside the spheres, respectively. The errors in the total energy of the studied molecules attributable to these numerical integrations, necessarily approximate, are of the order of 1.0 kJ/mol. Our accumulated experience in IQA calculations makes us believe that the accuracy achieved in each interatomic interaction is even higher. Despite this issue regarding the full numerical accuracy of our integrations, once the computational conditions of a given calculation have been chosen the conver-

Figure 24: Staggered BH_3NH_3, eclipsed BH_3NH_3, N_5^+, and Li_9H_9 molecules.



gence of the bipolar expansion (the exact benchmark) is ordinarily well below the 1 kJ/mol barrier for the xc contributions.

5.4 RESULTS AND DISCUSSION

The more representative results regarding the approximate V_{xc}^{AB} values, as well as their errors for the systems listed in Section 5.3 are gathered in tables 8-11 and figures 28-32. We can see in table 8, where the V_{xc}^{AB}'s for the HB systems of figure 23 are collected, that the full multipolar approximation $(V_{xc}^{AB})_{lr}$ (equation (5–15)) fails miserably for all intramolecular A—H pairs (A=N,O,F). Surprisingly, the crude (lr,cdq) approximation gives xc interactions with relative errors of about 10% or smaller for the intramolecular directly bonded atoms. Regarding the intermolecular interactions, the xc energy between the two atoms involved in the HB is well represented by $(V_{xc}^{AB})_{lr}$, with differences with respect to the exact values smaller than 0.3 kJ/mol in all the cases. We note again that the (lr,cdq) values differ only by 0.3-0.6 kJ/mol from the exact ones, confirming that the multipolar expansion for these interactions is practically converged at this level of calculation. Intermolecular A—H and H—H xc energies other than the above ones are given by the lr approximation with errors smaller than 0.1 and 0.01 kJ/mol, respectively. For the intermolecular H—H energies, the same is true in the (lr,cdq) approximation. However, the xc interaction between the A atom of the proton donor (PD) and the B atom of the proton acceptor (PA) molecule is predicted with errors as large as 1.4 kJ/mol ($FH\cdots NH_3$) when the (lr,cdq) approximation is used, which clearly indicates that multipolar interactions higher than the charge-quadrupole ones included

Figure 25: Molecules derived from saturated hydrocarbons by substituting C or H atoms by Be, B, N, O, F atoms, plus the benzene molecule.

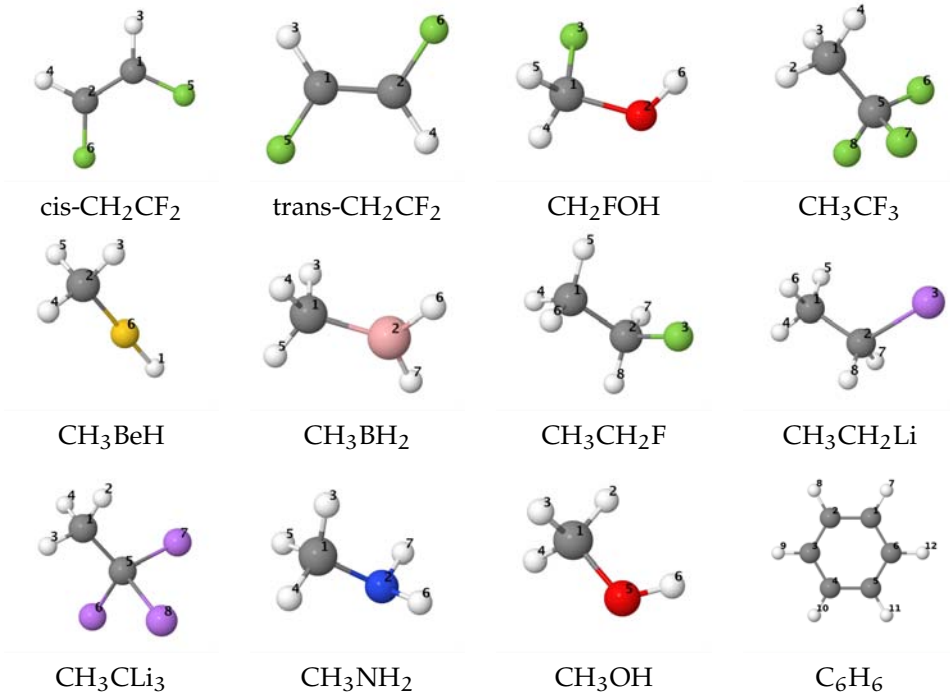


Figure 26: C_nH_{2n+2} $(n = 2-5)$ saturated hydrocarbons.

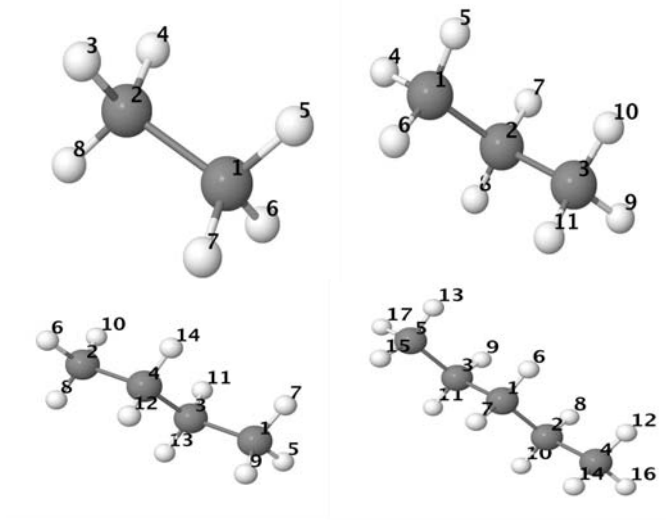
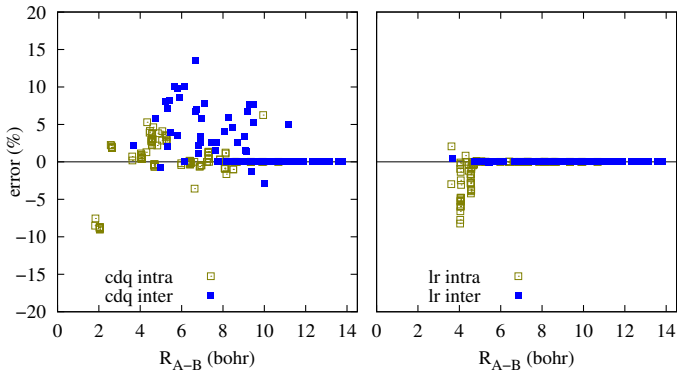


Table 8: xc interaction energies $\geqslant 0.1$ kJ/mol for the HB dimers of figure 23.

A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
H_2O−H_2O				**HF−NH_3**			
O_1−H_2	-478.59	1.29×10^{4}	-437.77	F_1−H_2	-280.40	3.79×10^{3}	-259.76
O_1−H_3	-569.61	-8.45×10^{5}	-514.46	F_1−N_3	-40.44	-41.74	-41.85
O_1−O_4	-16.40	-16.96	-16.94	F_1−H_4	-0.42	-0.42	-0.42
O_1−H_5	-0.11	-0.10	-0.11	F_1−H_5	-0.42	-0.42	-0.42
H_2−H_3	-1.94	-1.44	-1.84	F_1−H_6	-0.42	-0.42	-0.42
H_2−O_4	-24.73	-25.45	-25.36	H_2−N_3	-47.99	-48.50	-48.23
H_2−H_5	-0.17	-0.17	-0.17	H_2−H_4	-0.34	-0.34	-0.34
H_3−O_4	-0.19	-0.18	-0.18	H_2−H_5	-0.34	-0.34	-0.34
O_4−H_5	-541.14	-7.16×10^{5}	-488.52	H_2−H_6	-0.34	-0.34	-0.34
H_5−H_6	-2.19	-4.27	-2.03	N_3−H_4	-758.46	5.38×10^{5}	-688.30
H_2O−NH_3				N_3−H_5	-758.46	5.38×10^{5}	-688.30
O_1−H_2	-575.13	-1.21×10^{7}	-519.23	N_3−H_6	-758.40	5.58×10^{5}	-688.24
O_1−H_3	-460.75	-5.35×10^{4}	-421.84	H_4−H_5	-5.34	-12.72	-5.12
O_1−N_4	-20.16	-20.71	-20.70	H_4−H_6	-5.34	-12.71	-5.12
O_1−H_5	-0.19	-0.19	-0.19	H_5−H_6	-5.34	-12.71	-5.12
O_1−H_6	-0.20	-0.20	-0.20	**NH_3−H_2O**			
H_2−H_3	-1.88	-1.00	-1.80	N_1−H_2	-731.84	4.54×10^{6}	-668.43
H_2−N_4	-0.24	-0.24	-0.24	N_1−H_3	-779.59	1.37×10^{6}	-710.26
H_3−N_4	-31.54	-32.23	-32.07	N_1−O_5	-8.33	-8.63	-8.60
H_3−H_5	-0.27	-0.27	-0.27	H_2−H_3	-5.26	-6.18	-5.13
H_3−H_6	-0.28	-0.28	-0.28	H_2−O_5	-16.83	-17.59	-17.37
N_4−H_5	-764.79	6.93×10^{5}	-694.12	H_2−H_6	-0.10	-0.10	-0.10
N_4−H_6	-767.36	5.65×10^{5}	-697.06	H_3−H_4	-6.08	-30.52	-5.85
H_5−H_6	-5.60	-17.52	-5.37	H_3−O_5	-0.20	-0.20	-0.20
H_6−H_7	-5.66	-18.08	-5.43	O_5−H_6	-550.78	-8.08×10^{5}	-496.97
FHF^-				H_6−H_7	-2.27	-4.67	-2.11
F_1−F_2	-92.79	-100.02	-96.55	**NH_3−NH_3**			
F_1−H_3	-163.82	6.80×10^{4}	-159.19	N_1−H_2	-781.96	7.75×10^{5}	-712.71
HF−H_2O				N_1−H_4	-719.70	-1.22×10^{5}	-657.38
F_1−H_2	-298.63	-844.89	-275.72	N_1−N_5	-9.02	-9.29	-9.26
F_1−O_3	-28.64	-29.82	-29.79	H_2−H_3	-6.16	-35.32	-5.93
F_1−H_4	-0.23	-0.22	-0.22	H_2−H_4	-5.16	1.88	-5.04
H_2−O_3	-32.04	-32.78	-32.67	H_2−N_5	-0.22	-0.22	-0.22
H_2−H_4	-0.23	-0.23	-0.23	H_4−N_5	-20.79	-21.40	-21.28
O_3−H_4	-528.95	-5.93×10^{5}	-477.77	H_4−H_6	-0.24	-0.24	-0.24
H_4−H_5	-2.11	-4.02	-1.96	H_4−H_7	-0.20	-0.20	-0.20
HF−HF				N_5−H_6	-771.52	7.64×10^{5}	-702.05
F_1−H_2	-329.20	7048.78	-302.00	N_5−H_7	-770.75	6.85×10^{5}	-700.44
F_1−F_3	-17.02	-17.83	-17.82	H_6−H_7	-5.80	-22.66	-5.57
H_2−F_3	-16.71	-17.34	-17.32	H_7−H_8	-5.75	-21.54	-5.52
H_2−H_4	-0.11	-0.11	-0.11				
F_3−H_4	-356.83	-2.05×10^{5}	-320.36				
HF−N_2							
F_1−H_2	-349.90	-5.85×10^{3}	-319.67				
F_1−N_3	-9.19	-9.58	-9.57				
F_1−N_4	-0.28	-0.28	-0.28				
H_2−N_3	-11.15	-11.46	-11.46				
H_2−N_4	-0.46	-0.45	-0.45				
N_3−N_4	-2266.80	-8.61×10^{5}	-2479.99				

Figure 27: Phenol dimer.



in this approximation are required to represent this type of interaction with accuracy.

The relative errors of the A−B, A−H (A,B=N,O,F), and H−H *xc* interaction energies for all the intra and intermolecular pairs of the HB systems are represented in figure 28. We observe in figure 28 that cd and cdq intramolecular H−H energies are, in general, more accurate than the ones for the A−H interactions, which is clearly due to the 1−3 (1−2) character of all the intramolecular H−H (A−H) pairs. It is also striking that, with a couple of exceptions, cdq relative errors are negative whereas the contrary happens with the cd approximation. Moreover, as previously commented, only a single lr relative error appears in the figure, the remaining ones having errors greater than 20%. Regarding the intermolecular *xc* energies we observe in right figure 28 the progressive decreasing of relative errors in passing from cd to cdq, and from cdq to lr.

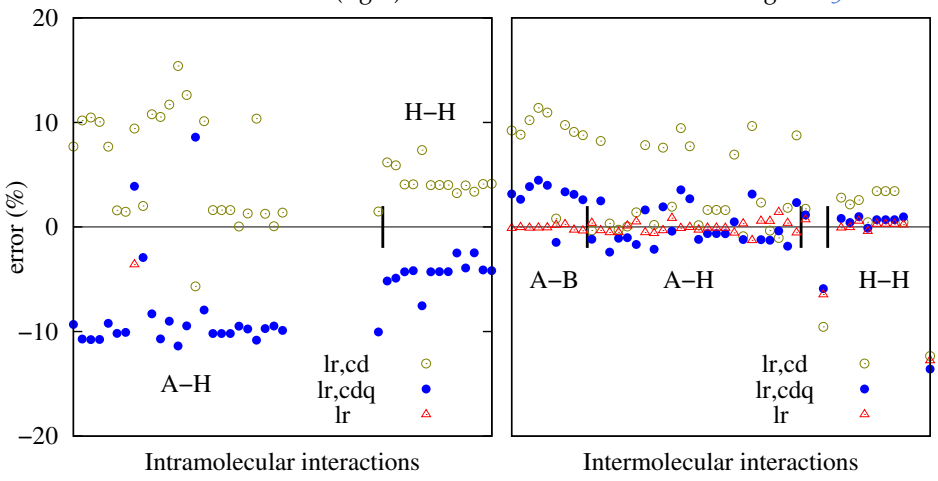
The discussion for the systems in figure 24 runs parallel to that of the HB dimers. The $(V_{xc}^{AB})_{lr}$ value for the B−N pair in eclipsed and staggered BH_3−NH_3 has no sense. Similarly, the lr *xc* interaction between the directly bonded (i.e. 1−2) B−H and N−H pairs is quite absurd. Not only that, but also the $(V_{xc}^{AB})_{lr}$'s for the 1−3 pairs H_3−H_4 and H_6−H_7 are several orders of magnitude greater than the exact values. Contrarily to this, the cdq approximation works relatively well for the B−N pair and the 1−2 B−H and N−H pairs (relative error $< 5\%$). The H_3−H_4 interaction is also extremely well reproduced by this approximation (error $< 0.2\%$), whereas the H_6−H_7 is slightly worse (error $\sim 4\%$). The *xc* interaction between the B atom and a H atom of the NH_3 unit (B_1−H_6) is fairly accurate in both the lr and cdq approximations. This is not so with the symmetric interaction N_2−H_3, with errors about 4−7% in both cases.

In the N_5^+ molecule, the *xc* energy for the 1−2 pairs is again badly represented by the lr approximation but is reasonable in the cdq approach, particularly for N_1−N_2. The cdq and lr values for the 1−3 N_1−N_4 interaction differs from the exact value by about 0.9 and 0.3 kJ/mol, respectively. The error in the

Table 9: Representative xc interaction energies (kJ/mol) for the systems of figure 24. In Li_9H_9 the directly bonded Li and H atoms are signalled with (1).

A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
eclipsed BH_3−NH_3				**Li_9H_9**			
B_1−N_2	-164.57	3524.65	-172.41	Li_1-Li_2	-0.30	-0.31	-0.31
B_1−H_3	-355.03	8.40×10^{6}	-334.69	Li_1-Li_6	-0.71	-0.72	-0.72
B_1−H_6	-1.15	-1.14	-1.14	Li_1-$H_{10}(1)$	-17.22	-318.57	-17.75
N_2−H_3	-60.32	-62.95	-65.25	Li_1-$H_{11}(1)$	-25.92	42.71	-26.24
N_2−H_6	-723.06	-8.05×10^{7}	-659.17	Li_1-H_{15}	-0.47	-0.49	-0.49
H_3−H_4	-67.52	-339.26	-67.39	Li_2-Li_3	-0.70	-0.93	-0.71
H_3−H_6	-2.58	-4.71	-2.55	Li_2-Li_4	-0.14	-0.14	-0.14
H_3−H_7	-0.69	-0.68	-0.69	Li_2-Li_6	-0.84	-0.87	-0.85
H_6−H_7	-4.30	-865.54	-4.12	Li_2-$H_{10}(1)$	-26.95	331.98	-27.12
staggered BH_3−NH_3				Li_2-$H_{11}(1)$	-18.34	59.37	-18.74
B_1−N_2	-172.72	3080.97	-181.51	Li_2-H_{12}	-0.47	-0.50	-0.49
B_1−H_3	-351.06	7.76×10^{6}	-331.51	Li_2-H_{13}	-0.09	-0.10	-0.09
B_1−H_6	-1.21	-1.99	-1.20	Li_2-$H_{15}(1)$	-36.35	1.50×10^{4}	-36.27
N_2−H_3	-64.56	-71.22	-69.56	Li_2-H_{17}	-0.17	-0.17	-0.17
N_2−H_6	-721.25	-6.056×10^{8}	-659.38	Li_6-Li_7	-0.31	-0.31	-0.31
H_3−H_4	-66.25	-296.87	-66.17	Li_6-H_{10}	-0.84	-0.88	-0.88
H_3−H_6	-1.06	-1.17	-1.05	Li_6-$H_{11}(1)$	-36.95	36.59	-36.67
H_3−H_7	-1.62	-1.61	-1.61	Li_6-H_{12}	-0.19	-0.19	-0.19
H_6−H_7	-4.28	-472.57	-4.12	Li_6-$H_{15}(1)$	-45.11	-1.21×10^{3}	-44.54
N_5^+				Li_6-H_{16}	-0.21	-0.21	-0.21
N_1−N_2	-1019.33	-1.70×10^{5}	-1056.45	H_{10}-H_{11}	-18.30	-20.78	-19.85
N_1−N_4	-47.28	-48.48	-48.20	H_{10}-H_{15}	-29.91	-38.97	-30.92
N_2−N_3	-52.15	-62.35	-55.59	H_{11}-H_{12}	-24.03	-36.36	-24.76
N_2−N_4	-2066.45	2.90×10^{7}	-2205.41	H_{11}-H_{13}	-0.70	-0.71	-0.70
N_2−N_5	-7.78	-7.77	-7.77	H_{11}-H_{15}	-25.73	-49.73	-26.56
N_4−N_5	-3.26	-3.27	-3.27	H_{11}-H_{17}	-0.28	-0.28	-0.28
				H_{15}-H_{16}	-1.48	-1.47	-1.44
				H_{15}-H_{17}	-0.19	-0.19	-0.19

Figure 28: Relative errors, $[(V_{xc}^{AB})_{method} - (V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}| \times 100$, of the intra- (left) and intermolecular (right) interactions of the molecules in figure 23.



other 1−3 interaction (N_2−N_3) in considerably higher in both approaches. Finally, the lr and cdq *xc* energies for the 1−4 N_2−N_5 and 1−5 N_4−N_5 pairs are practically the same and coincident with the exact value. This result highlights two important facts: i) the atomic basins of N_2 and N_5 (or N_4 and N_5) atoms fullfil almost exactly the non-overlapping criterion displayed in figure 22, and ii) the multipolar series (5–15) converges very quickly in this particular case.

Finally, the results for Li_9−H_9 reinforce what was said in the above three paragraphs. The full lr expansion fails completely in predicting *xc* interaction energies for 1−2 Li−H pairs, while the cdq values are pretty accurate. All Li−Li *xc* energies are well represented in the lr and cdq approximations, with the exception of the lr Li_2−Li_3 interaction. This is probably related with the almost spherical character of Li atomic basins. According to this, the 1−3 Li−H lr *xc* energies and, more importantly, the 1−3 H−H *xc* energies are less accurately computed due to the far from the spherical character of H atomic basins. This is exacerbated in the lr approximation, where higher angular number l values are involved (see equation (5–17)).

The *xc* pair interaction energies of the systems in figure 25 are collected in tables 10 and 11, and the relative errors of the cd, cdq and lr approximate values displayed in figure 29, for the 1−2 (left-top), 1−3 (right-top), and 1−4 (bottom) pairs, respectively. Virtually all of the above comments also apply here: the 1−2 *xc* interactions can not be represented at all by using the full lr expansion. However, they are given with reasonable accuracy by the cdq approximation. The 1−n ($n > 2$) interactions are gradually better reproduced as n increases in both the lr and cdq approximations. It is very satisfactory to check that the cdq approach, a severe truncation of the full *mp* expansion, is perfectly suited to simulate the *xc* interaction between pairs of atoms beyond the directly bonded ones. Even in typically covalent molecules like benzene all the cdq C−C *xc*

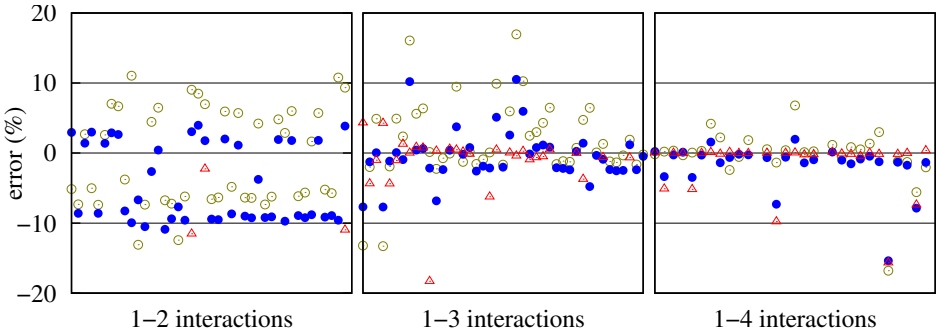
Table 10: Representative xc interaction energies (kJ/mol) for the systems of figure 25.

A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
	cis−$C_2H_2F_2$				CH_3BeH		
C_1−C_2	-1311.48	-3.82×10^5	-1351.45	H_1−C_2	-34.84	-32.52	-32.62
C_1−H_3	-811.14	-1.20×10^4	-746.88	H_1−H_3	-0.75	-0.74	-0.75
C_1−H_4	-17.00	-15.11	-15.79	H_1−Be_6	-170.49	2.43×10^4	-159.81
C_1−F_5	-600.29	-5.29×10^3	-608.87	C_2−H_3	-845.84	6.99×10^5	-765.34
C_1−F_6	-53.21	-54.86	-52.56	C_2−Be_6	-168.87	-824.46	-164.53
H_3−H_4	-1.40	-1.40	-1.40	H_3−H_4	-17.84	-30.31	-17.43
H_3−F_5	-24.53	-24.79	-24.54	H_3−Be_6	-2.46	-2.45	-2.47
H_3−F_6	-3.64	-3.70	-3.52		CH_3BH_2		
F_5−F_6	-4.61	-4.60	-4.60	C_1−B_2	-361.03	-496.34	-362.60
	trans−$C_2H_2F_2$			C_1−H_3	-826.73	2.48×10^5	-745.50
C_1−C_2	-1304.91	-3.87×10^5	-1345.33	C_1−H_4	-834.27	6.39×10^4	-762.61
C_1−H_3	-811.92	-540.92	-747.51	C_1−H_6	-66.37	-68.63	-68.96
C_1−H_4	-16.93	-15.04	-15.72	B_2−H_3	-6.72	-6.70	-6.71
C_1−F_5	-603.33	-5.67×10^3	-611.80	B_2−H_4	-4.08	-4.11	-4.11
C_1−F_6	-54.31	-56.04	-53.69	B_2−H_6	-376.42	-1.85×10^4	-349.54
H_3−H_4	-1.37	-1.38	-1.38	H_3−H_4	-17.11	-25.73	-16.68
H_3−F_5	-24.82	-25.10	-24.84	H_3−H_6	-2.04	-2.04	-2.04
H_3−F_6	-3.66	-3.72	-3.53	H_4−H_5	-15.65	-21.75	-15.36
F_5−F_6	-4.71	-4.69	-4.69	H_4−H_6	-2.82	-2.81	-2.81
	CH_2FOH			H_4−H_7	-2.11	-2.10	-1.72
C_1−O_2	-595.20	1.19×10^3	-613.00	H_6−H_7	-66.83	-69.51	-65.42
C_1−F_3	-515.50	80.74	-529.59		CH_3CH_2F		
C_1−H_4	-784.72	-1.93×10^4	-724.78	C_1−C_2	-776.25	-808.13	-790.31
C_1−H_6	-3.15	-3.08	-3.12	C_1−F_3	-34.45	-36.53	-36.63
O_2−F_3	-84.54	-94.13	-94.16	C_1−H_4	-818.55	5.40×10^4	-748.07
O_2−H_4	-29.48	-29.36	-29.61	C_1−H_5	-820.78	7.25×10^4	-749.52
O_2−H_6	-511.47	-2.86×10^5	-465.17	C_1−H_7	-14.33	-14.45	-14.31
F_3−H_4	-31.06	-31.02	-31.26	C_2−F_3	-546.94	10.18	-558.24
F_3−H_6	-5.11	-5.18	-5.19	C_2−H_4	-17.73	-17.99	-17.87
H_4−H_5	-11.89	-13.77	-11.64	C_2−H_5	-15.64	-15.90	-15.82
H_4−H_6	-0.47	-0.47	-0.47	C_2−H_7	-804.65	2.30×10^5	-740.40
	CH_3CF_3			F_3−H_4	-2.92	-2.90	-2.90
C_1−H_2	-818.30	8.69×10^4	-746.71	F_3−H_5	-3.47	-3.54	-3.54
C_1−C_5	-735.49	-846.24	-758.72	F_3−H_7	-31.02	-31.13	-31.29
C_1−F_6	-31.84	-33.40	-33.56	H_4−H_5	-15.48	-21.10	-15.16
H_2−H_3	-13.27	-16.44	-13.01	H_4−H_7	-1.07	-1.06	-1.06
H_2−C_5	-16.37	-16.79	-16.80	H_5−H_6	-15.14	-20.32	-14.82
H_2−F_6	-2.55	-2.53	-2.53	H_5−H_7	-1.18	-1.17	-1.17
H_2−F_7	-3.45	-3.53	-3.21	H_5−H_8	-2.08	-2.08	-4.90
C_5−F_6	-469.52	-792.76	-488.98	H_7−H_8	-13.54	-17.74	-13.22
F_6−F_7	-83.69	-93.89	-93.53				

Table 11: Representative xc interaction energies (kJ/mol) for the systems of figure 25 (cont).

A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	A − B	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
CH_3CH_2Li				**CH_3NH_2**			
C_1-C_2	-789.95	9.49×10^3	-798.97	C_1-N_2	-740.43	1.21×10^3	-753.96
C_1-Li_3	-1.66	-1.67	-1.67	C_1-H_3	-806.63	1.18×10^4	-740.54
C_1-H_4	-810.07	1.51×10^5	-743.11	C_1-H_4	-809.44	9.59×10^4	-741.04
C_1-H_5	-815.31	2.79×10^5	-746.34	C_1-H_6	-6.68	-7.10	-6.60
C_1-H_7	-21.37	-22.48	-21.68	N_2-H_3	-29.22	-29.36	-28.96
C_2-Li_3	-100.43	-554.91	-96.79	N_2-H_4	-27.42	-27.66	-27.76
C_2-H_4	-19.31	-36.64	-18.43	N_2-H_6	-770.22	1.45×10^6	-707.83
C_2-H_5	-18.60	-13.76	-18.54	H_3-H_4	-16.11	-21.94	-15.75
C_2-H_7	-838.53	3.78×10^6	-767.82	H_3-H_6	-0.59	-0.58	-0.58
Li_3-H_4	-0.57	-0.57	-0.57	H_4-H_5	-15.88	-22.49	-15.48
Li_3-H_5	-0.16	-0.16	-0.16	H_4-H_6	-0.62	-0.61	-0.61
Li_3-H_7	-1.98	-1.98	-1.96	H_4-H_7	-1.91	-1.90	-1.77
H_4-H_5	-18.94	-36.17	-18.51	H_6-H_7	-5.87	-16.69	-5.67
H_4-H_7	-1.59	-1.57	-1.57	**CH_3OH**			
H_5-H_6	-19.74	-36.04	-19.25	C_1-O_2	-638.41	-187.46	-650.07
H_5-H_7	-1.26	-1.25	-1.25	C_1-F_3	-808.68	-6.14×10^3	-740.99
H_5-H_8	-2.65	-2.64	-2.64	C_1-H_4	-804.40	2.53×10^4	-738.45
H_7-H_8	-23.12	-97.48	-22.57	C_1-H_6	-3.79	-3.81	-3.78
CH_3CLi_3				O_2-F_3	-36.31	-36.22	-36.71
C_1-H_2	-806.90	4.89×10^5	-739.52	O_2-H_4	-29.92	-29.73	-30.14
C_1-C_5	-823.71	2.89×10^4	-839.98	O_2-H_6	-553.82	-3.09×10^5	-503.97
C_1-Li_6	-2.38	-2.42	-2.41	F_3-H_4	-15.27	-21.13	-14.92
H_2-H_3	-20.16	-63.55	-19.70	F_3-H_6	-1.32	-1.31	-1.32
H_2-C_5	-25.16	-33.67	-25.04	H_4-H_5	-14.69	-18.89	-14.35
H_2-Li_6	-0.65	-0.64	-0.64	H_4-H_6	-0.50	-0.50	-0.50
H_2-Li_7	-0.43	-0.43	-0.38	**C_6H_6**			
C_5-Li_6	-135.99	-1.03×10^5	-123.94	C_1-C_2	-1043.99	-2.45×10^6	-1065.06
Li_6-Li_7	-1.34	-3.02	-1.29	C_1-C_3	-23.85	-26.79	-24.50
				C_1-C_4	-22.49	-23.40	-23.40
				C_1-H_7	-819.10	1.90×10^7	-756.10
				C_1-H_8	-16.89	-18.55	-17.07
				C_1-H_9	-1.73	-1.73	-1.73
				C_1-H_{10}	-0.92	-0.93	-0.93
				H_7-H_8	-1.84	-1.86	-1.84
				H_7-H_9	-0.21	-0.21	-0.21
				H_7-H_{10}	-0.05	-0.05	-0.05

Figure 29: Relative errors, $[(V_{xc}^{AB})_{method} - (V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}| \times 100$, for the $1-2$, $1-3$, and $1-4$ interactions of the molecules represented in figure 25. Empty circles, bold circles, and triangles stand for (lr,cd), (lr,cdq), and lr calculations, respectively.



interaction energies reproduce very well the exact values. As we can see in figure 29 most of the $1-2$ interactions have relative cdq errors $\leqslant$ 10%. This improves for the $1-3$ and $1-4$ interactions.

A summary of our results for the saturated hydrocarbons C_nH_{2n+2} ($n = 2-5$) is presented in graphical form in figure 30. We find the surprising result that all $1-2$ C–C cdq interactions are predicted with errors $\leqslant$ 2% while the cdq energies between the more distant $1-3$ C–C pairs have errors about 5–6%. Nevertheless, the interactions between even more distant C–C pairs turn again to be calculated quite accurately (errors $<$ 1%) in the cdq approximation. The lr approximation fails completely to predict the $1-2$ C–C interactions, but yields negligible errors for the $1-3$ *xc* interaction energies. With regard to the C–H interactions, the situation is the opposite of that found for the C–C pairs: 1-2 C–H cdq errors are about 9-10% (except in ethane where the error is unusually large (54%)) whereas all except two of the 1-3 C–H cdq errors are $<$ 1%. For these two exceptions the error is not too large ($\sim$ 1.3%). We observe in figure 30 that the cdq approximation improves considerably the cd results, giving 1-3 C–H interaction energies almost as accurate as the lr ones. Another surprising result in these systems concerns the cdq $1-3$ H–H interactions: Contrary to what happens almost systematically, the cdq results are worse than the cd ones, albeit the relative errors in both approximations are acceptable ($\sim 2-3$%).

We have considered staggered ethane as a representative example to analyze the correlation effects on the V_{xc}^{AB} energies. Our results for five representative AB pairs of this molecule are collected in table 12. Correlation decreases (increases) the magnitude of the C-C interaction (H-H interactions), changes very little V_{xc}^{CH} when C belongs to a CH_3 group and H to the other, and enhances the intra-group V_{xc}^{CH} energies. The discussion of the above paragraph for the RHF results is still approximately valid for the correlated calculation. With the exception of inter-group H-H interactions, the cdq xc energies are closer to the exact V_{xc}^{AB} values than their full multipolar approximations $(V_{xc}^{AB})_{lr}$. As in

Figure 30: Relative errors, $[(V_{xc}^{AB})_{method} - (V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}| \times 100$, for the $1-2$ (left) and $1-3$ (right) interactions of the molecules represented in figure 26.

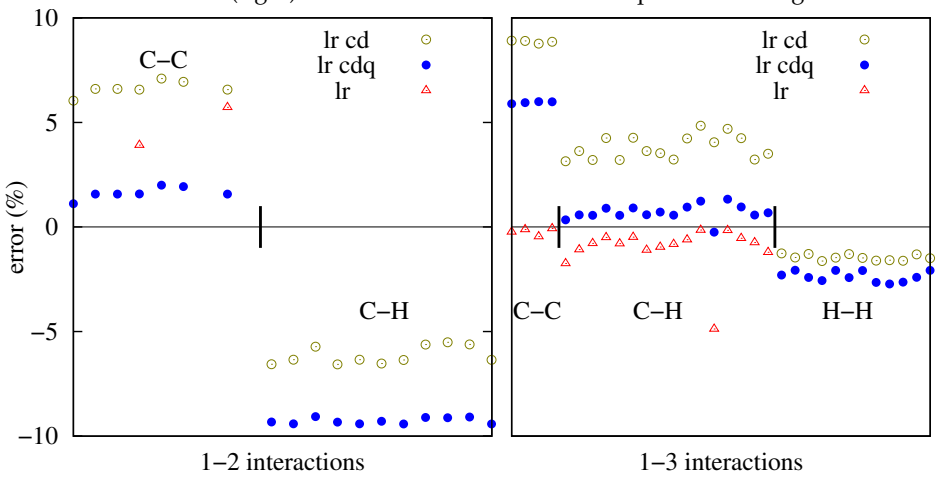


Table 12: Comparison of RHF and CAS[14,14] xc energies (kJ/mol) for staggered C_2H_6.

	RHF			CAS[14,14]		
	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
C_1-C_2	-781.87	-870.42	-790.68	-625.88	-723.66	-639.75
C_1-H_5	-821.98	8.93×10^4	-533.05	-686.06	8.17×10^4	-618.64
H_3-H_4	-16.75	-24.11	-16.38	-25.06	-32.82	-24.75
C_1-H_3	-16.21	-16.55	-16.26	-16.62	-16.75	-16.55
H_3-H_6	-2.32	-2.32	-2.32	-3.10	-3.10	-3.10

many of the HF $1-2$ interactions, the lr approximation fails to predict even a reasonable value for the C_1-H_5 *xc* energy.

The different behavior of the lr and cdq approximations can be further illustrated with the case of the phenol dimer (figure 27). For this system, the relative errors *versus* the interatomic distance R_{A-B} in these two approximations are plotted in figure 31, both for intra-molecular and inter-molecular atomic pairs. Only two points, associated to intra-molecular interactions, have a relative error (absolute value) $\geqslant 20\%$ in the cdq calculation, while the error for all the lr points with $R_{A-B} < 2.64$ (most of them associated to intra-molecular pairs) is larger than 20%. However, for $R_{A-B} > 5.0$ the lr approximation gives quite accurate *xc* interaction energies for all the pairs, whereas cdq errors are still important.

However, there is a general problem of the lr approximation that deserves to be commented: equation (5–15) does not necessary converges to the exact *xc* interaction for large $l_1 m_1$ and $l_1 m_1$ values. This fact is illustrated in figure 32, where the *xc* energies for some of the atomic pairs of the molecules in figure 25 are represented *versus* $l_1^{max} + l_2^{max}$. The lr approximation suffers a systematic error in the C_1-F_6 interaction of cis-CH_2CF_2 and trans-CH_2CF_2 molecules,

Figure 31: Relative error of the cdq (left) and lr (right) calculations for the phenol-dimer ($C_6H_5OH\cdots C_6H_5OH$). Only two (2) points are out of the ordinate scale in the cdq calculation, while all the points (26) with $R_{A-B} < 2.64$ bohr are out of the ordinate scale in the lr calculation.

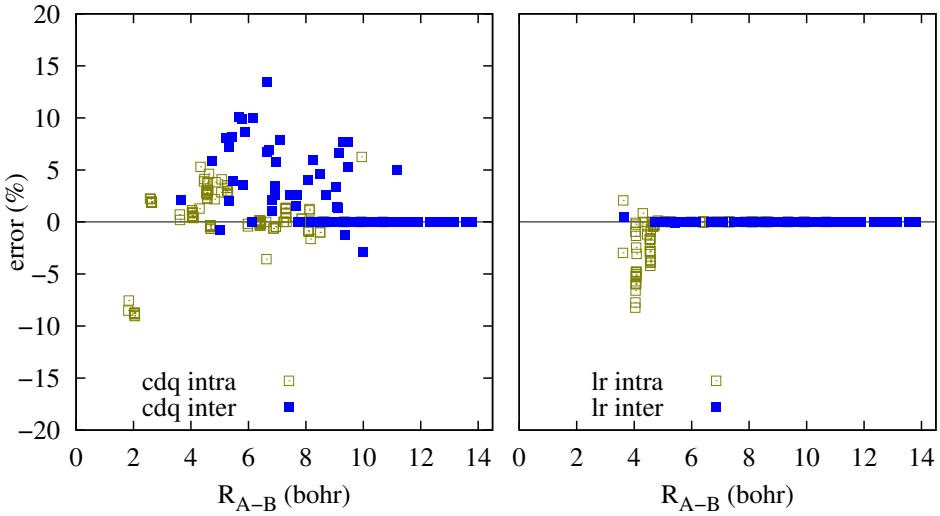


Figure 32: Convergence of the $A-B$ interactions indicated in the figure. The pairs i, j correspond to the labels of figure 25

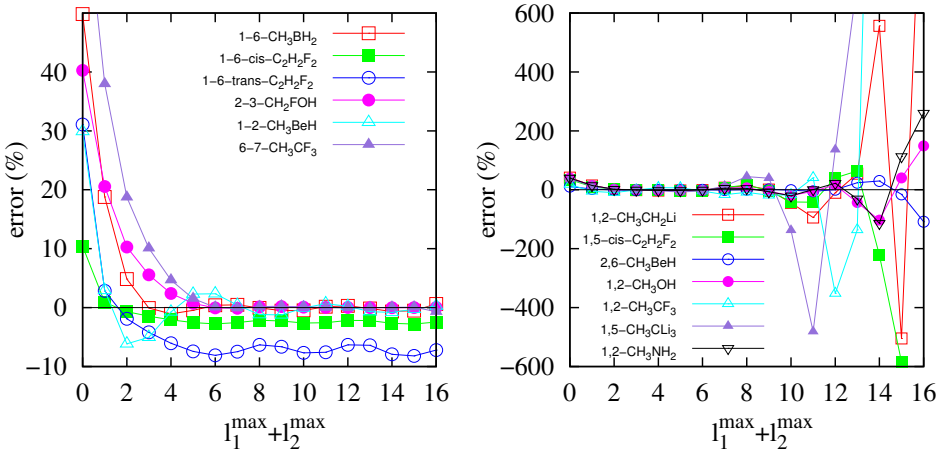
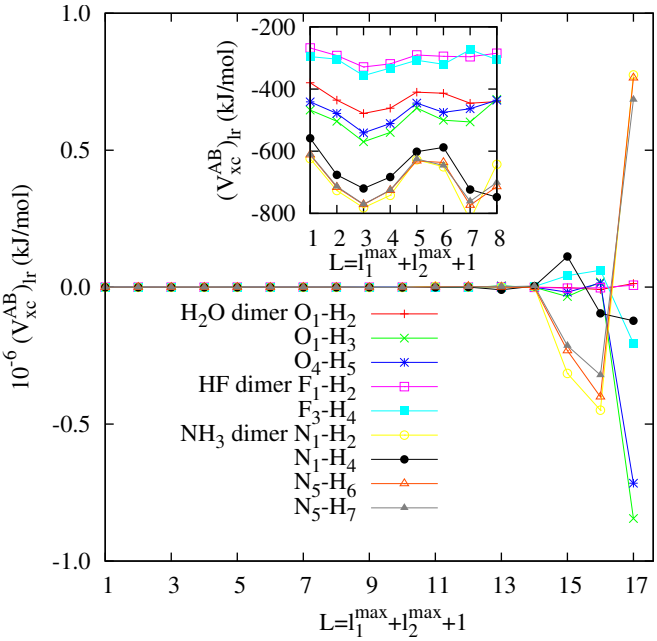


Figure 33: Convergence of $1-2$ interactions in the $H_2O{-}H_2O$, HF-HF and $NH_3{-}NH_3$ dimers.

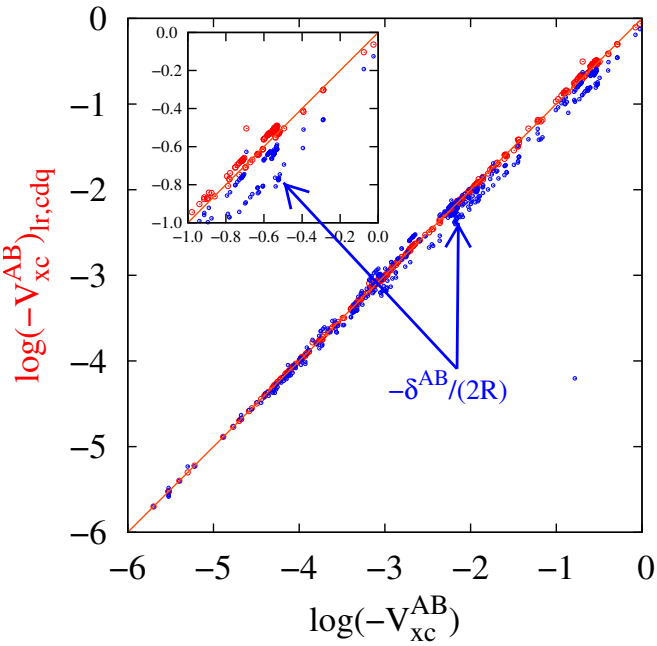


regardless the value of $l_1^{max}+l_2^{max}$. In the case of trans-CH_2CF_2 the $C_1{-}F_6$ *xc* interaction energy shows an oscillating behavior around an (erroneous) mean value. This pattern has been also observed in other cases. Contrarily, as we have repeatedly said in this section, catastrophic lr interactions (see, for instance figure 32) are still reasonable provided that the sum l_1+l_2 is interrupted at a value approximately in the interval $2 \leqslant l_1+l_2 \leqslant 6$.

This simple analysis shows that conclusions on the convergence of the *mp* expansion drawn exclusively from limited $L = l_1^{max}+l_2^{max}+1$ data cannot be trusted. Figure 33, where the $(V_{xc}^{AB})_{lr}$ energies for some $1-2$ interactions of the $H_2O{-}H_2O$, HF-HF and $NH_3{-}NH_3$ dimers are shown, illustrates this fact in a crystal clear way. Cutting the *mp* expansion at $L \leqslant 8$ the oscillatory behavior of figure 32 would have been found indeed, but no catastrophe for larger L values would have been predicted. However, for $L > 8$ the *mp* expansion progressively deteriorates and for high L values the $(V_{xc}^{AB})_{lr}$ energies diverge. Again, the cdq approximation works fairly well, with all the $1-2$ interactions in figure 33 predicted with relative errors smaller than 3% except $F_1{-}H_2$ (3.2%) and $F_3{-}H_4$ (5.2%).

Summarizing, we have shown that the multipole series for the interatomic xc energies is conditionally convergent, and that the computational burden of the quasi-exact calculation of V_{xc} when general QTAIM domains are used may be ameliorated by retaining up to quadrupole-quadrupole terms. With this approximation, reasonable errors are obtained in the medium- to long-distance

Figure 34: Comparison of all the exact V_{xc}^{AB} values considered in this work to the monopole-monopole (small circles) and cdq approximations (large circles) in a logarithmic scale.



range, sometimes even for directly bonded interactions. An overall image of the improvement of the cdq approximation over the cc (monopole-monopole) one can be grasped from figure 34 that condenses all our calculations that span a six orders of magnitude range for V_{xc}.

We have explored numerically the degree of fulfillment of the linear relation between V_{xc}^{AB} and $\delta^{AB}/(2R)$. Assuming $V_{xc}^{AB} \simeq -a[\delta^{AB}/(2R)] + b$, we have determined a and b for each molecule by fitting the computed values of V_{xc}^{AB} and δ^{AB} for every AB pair of this molecule to the above expression, verifying that b is always very small and a takes values relatively close to (but smaller than) 1.0. As representative examples, (a, b) for the H_2O-H_2O, NH_3-NH_3, and HF-HF dimers are $(a, b) = (0.8988, -0.0005)$, $(0.8673, 0.0000)$, and $(0.9150, -0.0010)$, respectively. This result opens another possible route to the approximate but much cheaper computation of the *xc* interaction in those cases where the exact calculation is prohibitive or very expensive. An extensive analysis of this correlation in anion-π interactions which corroborates the above statement has been carried out by Foroutan-Nejad *et. al.* [13, 48, 132]

5.5 CONCLUSIONS

We have shown that the interatomic exchange-correlation energies used in real space theories of chemical bonding, which measure the covalent contribution to a given interatomic interaction, can be approximated via a conventional multipole expansion. Rigorously, the series diverges when atoms are directly bonded, although it may be regarded asymptotically convergent. Truncation of the series up to $l_1 + l_2 = 2$ (including up to charge-quadrupole interactions) tends to provide results which are accurate to a few percent in 1-n, $n > 2$ interactions, and even to about 10% in many 1-2 directly bonded cases. In the $n > 2$ case the series converges in many cases, and including extra terms provides further accuracy. On the contrary, the consideration of larger l contributions in 1-2 interactions tends to seriously deteriorate the results. Since the computational burden needed to calculate the multipole series is considerably smaller than that of the exact bipolar expansion, our results may be important to estimate covalent interactions in those cases where exact integrations are not feasible. They can also be used to ameliorate the computational cost in IQA decompositions of large systems, where many expensive, but small long-range *xc* terms can now be safely approximated without loss of precision.

Part III

Molecular properties from bonding concepts

6 A CHEMICAL VIEW OF THE LOCALIZATION TENSOR

CONTENTS

While the modern theory of the insulating state shows that the conducting or insulating properties of a system can be extracted solely from ground state properties via the so-called localization tensor (LT), no chemical reading of this important quantity has ever been offered. Here, a remarkable link between the LT and bond orders as described by the delocalization indices (DIs) of chemical bonding theory is reported. This is achieved through a real space partition of the LT into intra- and interatomic contributions. We show that the convergence or divergence of the LT in the thermodynamic limit, which signals the insulating or conducting nature of an extended system, respectively, can be nailed down to DIs. This allows for the exploitation of traditional chemical intuition to identify essential and spectator atomic groups in determining electrical conductivity. The thermodynamic limit of the LT is controlled by the spatial decay rate of the interatomic DIs, exponential in insulators and power-law in conductors. Computational data in a few selected toy systems corroborate our results.

6.1 INTRODUCTION

As the technological demand of smart, functional, or taylored materials increases, so does the need for understanding the basic physics behind their sought-after properties. In many cases this search has led to explore the new dimension that the dependence of physical properties on size introduces at the nanoscale. For instance, the predicted demise of Moore's law [189] has stirred up the development of new quantum-mechanically operated devices like the single electron transistor [57]. Similarly, new fields such as molecular electronics have become hot topics producing thousands of specialized papers [40]. Despite much work, the building of new physical or chemical intuition that may guide future research beyond that coming from brute force case-by-case simulation has proven much more difficult. In crystal engineering, as an example, although the situation is now much better than 20 years ago, we are still far from mastering the rules to synthesize on-demand crystal structures [43].

Regarding electrical conductivity at the nanoscale, much work has been devoted in molecular electronics to quantitatively simulate electron transport in single-molecule junctions [188], and some rules regarding the factors that govern their conductivity have emerged. However, despite the efforts, no simple chemical rules linking molecular structure and molecular conductivity have been found todate [40]. Since, in the end, all newly developed nanodevices depend on the chemical synthesis of taylored molecular fragments, we believe that finding simple chemical indicators of facile electronic transport or conductivity is an important goal with possibly major outcomes.

A guiding principle in this quest may be taken from the naïve chemical association between conductivity and electron localization and delocalization. Key concepts in chemistry like conjugation, resonance, aromaticity, etc, are nothing but different incarnations of electron localizability. However, standard

approaches coming from the theory of chemical bonding (TCB) are almost inevitably linked to the one-particle molecular orbital (MO) theory [69], and molecular conductivity tends to be interpreted in terms of excitation gaps, i.e. HOMO-LUMO energetic differences, instead of as a ground state property that could be transformed into the sought conductivity indicators. Fortunately, a new paradigm in TCB has emerged in the last few decades [10] that defines (and explores) chemical objects in real space from orbital invariant densities (or density matrices). These techniques, collectively known as quantum chemical topology (QCT) [154], analyze the wave function of a system, and use meaningful fields to partition the physical space into regions or domains associated to: atoms, through the one particle density in the quantum theory of atoms in molecules [10] (QTAIM); cores, lone, and bonding pairs, through the electron localization function[174] (ELF) or the electron localizability indicator [90] (ELI), etc. Once the real space objects are defined, indicators are obtained at well-defined points, usually the critical points of the defining field, and the global expectation values of operators are divided into domain contributions. This allows, for instance, for a rigorous real space partitioning of the energy into intra- and interatomic components (the interacting quantum atoms approach [21], (IQA) much in the spirit of the atomistic ansatz. QCT, well-known in quantum chemistry, is slowly entering condensed matter physics.

Early attempts that tried to link the conductivity features of a molecular system with its electron density failed [31]. This comes as no surprise, since conductivity leaves no simple scars on the density. Fortunately, QCT domain expectation values are based on physical observables, so QCT provides an open door to connect the physicist's and the chemist's intuitions, which tend to live in separated worlds. This is not easy to do in other TCB approaches. In this regard, a rigorous formalism coupling the insulating or conducting nature of an extended system with ground state properties exists [88]. Although not well known in the chemical literature, Kohn's theory of the insulating state does the job. It is electron (de)localization that explains conductivity, quantified by an object called the localization tensor (LT).

Thanks to QCT and its rigorous partitioning of quantum mechanical expectation values into atomic or functional group contributions, we find and explore here a remarkable bridge between the LT and the standard bond orders of chemistry, as defined in their real space manifestation known as delocalization indices (DIs). It is the rate at which bond orders decrease with distance that determines whether a system will or will not be conducting in the static thermodynamic limit. Since we can examine straightforwardly the behavior of DIs among atoms or functional groups in several dimensions, we expect our results to be useful in building new conductivity chemical rules.

6.2 THE MODERN THEORY OF THE INSULATING STATE AND THE ASSESSMENT OF CONDUCTIVITY VIA THE LOCALIZATION TENSOR

A seminal work by W. Kohn in 1964 [88] showed for the first time how the insulating nature of a system could be understood as a consequence of electron localization in the ground state, and not only from the properties of its excitation spectrum. However important, this line of reasoning remained largely unexplored until the end of the 1990's, when Resta revisited and generalized it [160, 161, 163]. As emphasized by this author [162], it is the organization of electrons in the ground state that renders a system insulating or conducting. A central object that quantifies Kohn's localization in an N electron system is the localization tensor, $\boldsymbol{\lambda}$ or LT. It is defined as the second cumulant moment, per electron, of the total electronic position operator $\hat{\mathbf{R}} = \sum_i^N \hat{\mathbf{r}}_i$.

$$\boldsymbol{\lambda} = \frac{1}{N}\left\{\langle\Psi|\hat{\mathbf{R}}\otimes\hat{\mathbf{R}}|\Psi\rangle - \langle\Psi|\hat{\mathbf{R}}|\Psi\rangle\otimes\langle\Psi|\hat{\mathbf{R}}|\Psi\rangle\right\}. \tag{6–1}$$

We will use in this work bold fonts to indicate vectors or tensors, depending on the context, and the $\otimes$ symbol for tensor or cartesian products. As an example, the cartesian components of the $\mathbf{r}\otimes\mathbf{r}$ tensor are $(\mathbf{r}\otimes\mathbf{r})_{\alpha\beta} = x_\alpha x_\beta$.

One of the most important results of Resta's reformulation lies in the link between the behavior of $\boldsymbol{\lambda}$ in the thermodynamic limit and electrical conductivity: the $\boldsymbol{\lambda}$ tensor, that measures the quadratic fluctuation of the polarization of the system, and that was initially used by Kudinov [95], has a well-defined thermodynamic limit, diverging for conductors while remaining finite for insulators.

We will just provide, for consistency, a few ideas that may guide the informed reader about the origin of such a unique property. It stems from the fluctuation-dissipation theorem [176], that allows to prove [162] first that

$$\boldsymbol{\lambda} = \frac{\hbar}{\pi e^2 N}\int_0^\infty d\omega\ \mathrm{Im}\boldsymbol{\alpha}(\omega), \tag{6–2}$$

where $\boldsymbol{\alpha}(\omega)$ is the frequency dependent linear polarizability tensor. From this, if periodic boundary conditions are imposed, it can also be proven that

$$\lambda_{\beta\gamma} = \delta_{\beta\gamma}\frac{\hbar V}{\pi e^2 N}\int_0^\infty d\omega\ \frac{\mathrm{Re}\sigma(\omega)}{\omega}, \tag{6–3}$$

σ being the frequency dependent electric conductivity. For conducting systems, with non-vanishing Reσ at zero frequency, the diagonal components of $\boldsymbol{\lambda}$ diverge. These diagonal values can also be understood as localization lengths [176], and they are related to the optical gap E_g by $\lambda_{\alpha\alpha} \leqslant \hbar^2/(2m_e E_g)$.

Simple manipulations, already put forward by Resta [162], allow to recast the LT in terms of the first order, $\rho(\mathbf{r}_1)$, and the second order, $\rho_2(\mathbf{r}_1,\mathbf{r}_2)$, spinless densities introduced in section 1.3.1.

Using the exchange-correlation density, the part of the pair density containing the quantum mechanical effects due to the antisymmetry of the wavefunction, $\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2) = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) - \rho_2(\mathbf{r}_1,\mathbf{r}_2)$, and defining the interparticle position vector $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$, $\boldsymbol{\lambda}$ may be written in an explicitly origin independent, symmetrical form. As we show in the ESI, after some algebraic manipulations,

$$\boldsymbol{\lambda} = \frac{1}{2N}\int d\mathbf{r}_1 d\mathbf{r}_2 \ (\mathbf{r}_{12}\otimes\mathbf{r}_{12})\,\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2). \tag{6–4}$$

6.3 A BRIDGE BETWEEN CONDUCTIVITY AND THE THEORY OF CHEMICAL BONDING

Being Chemistry the science of the interactions among electrons (or atoms, made from them and nuclei), it is not surprising that ρ_{xc}, that collects all non-classical behaviour in the pair-density, is emerging, slowly but steadily, as one of the pillars in the modern TCB. Similarly, having shown that electrical conductivity is related to how electrons localize or delocalize, it is also expectable that the former be related to chemical bonding measures of electron delocalization. Undoubtedly, the latter have a rather long history in TCB.

In the present context, Bader and Stephens [9] already proposed in 1974 that the interatomic integration of ρ_{xc} measures the number of pairs of electrons shared between two atomic regions, and named this quantity the delocalization index (DI),

$$\delta^{AB} = 2\int_A d\mathbf{r}_1 \int_B d\mathbf{r}_2 \ \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2). \tag{6–5}$$

Here A, B are the spatial regions associated to the two atoms (or fragments) under scrutiny. A similar A, A integral, the localization index, determines the number of localized electrons in a region. The DI provides the fluctuation of the electron population in the A, B regions, being also a real space generalization of the standard MO Wiberg-Mayer bond order [119, 190], which physicists still use, in their majority, in its even cruder Mulliken flavor. In energetic terms, the interatomic exchange-correlation energy,

$$E_{xc}^{AB} = \int_A d\mathbf{r}_1 \int_B d\mathbf{r}_2 \ \frac{\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)}{r_{12}} \tag{6–6}$$

has been shown to correspond to the covalent part of the interaction between the regions [109]. Besides these two direct links, a growing body of evidence is showing the relevance of ρ_{xc}-based indices in TCB [65], explaining facts as the nature of chemical interactions from DI profiles [65], or rationalizing stereolectronic effects [144]. As we are going to show, *it is the innocent concept of bond order, a must in every freshman chemistry course, that stores information about electrical conductivity*.

A couple of recent studies [61, 62], had already started to show that DIs encode information about the insulating or conducting nature of a system

through their spatial decay rate: in metals we find an algebraic oscillatory decline with interatomic $A - B$ distance, while in insulators their fall off is exponential. Consideration of strongly correlated cases [62], evidences that DIs are also suitable generalizations of the double occupation order parameter $\mathcal{D}$ used in Hubbard models to signal metal-insulator transitions, and that they reveal how mesomeric effects in alternant hydrocarbons are deeply linked to the oscillatory pattern that leads to conductivity in the thermodynamic limit. An increase in the electron correlation strength (by increasing the Hubbard U/t parameter or equivalently by substituting carbon by heavier elements) eventually destroys the oscillations, pointing toward an active effect of electron correlation in chemistry, e.g. to smaller mesomeric effects in the heavier analogues of alternant hydrocarbons. Even more importantly, the decay of these indices may be followed along specific bond chains, directions, or along a combination of both. One needs only choose appropriately the domains in the A, B pairs. This provides a quantitative tool in the discovery of low dimensional conductors.

6.3.1 The localization tensor in finite molecules

In the last few years, several works by Leininger, Evangelisti and coworkers [29, 44] have examined the role of $\boldsymbol{\lambda}$ in molecular instead of extended systems. To that end, these authors have preferred to use the total second cumulant, which they have called total position-spread tensor, $\boldsymbol{\Lambda}$ or TPS, and not the per electron quantity. Even a spin resolved version has also been studied [45].

It has been shown that the TPS is very sensitive to bond stretching, becoming large in the case of increased electron mobility. In simple diatomics, for instance, its parallel component is small at equilibrium, it increases as the interatomic distance is enlarged before achieving a maximum value close to the bond breaking region, and it decreases again towards the free atomic value at dissociation.

Although the TPS has been wellcome, adding to the battery of new chemical bonding indicators at hand, its global character partially limits its applicability. Its evolution in a possibly complex process will just average out the total response of the system, even though some very restricted atomic or bond resolution might be achieved by following a particular component or projection that isolates an important direction in space. In order to become a useful TCB descriptor, this barrier needs be overcome to understand the origin of convergence/divergence and the onset of conductivity as we approach the thermodynamic limit.

6.3.2 An atomic partition of the TPS

Being the expectation value of a two-electron operator, QCT offers an immediate solution to the problem: provided that a chemically meaningful division of the space exists, we can space partition $\boldsymbol{\lambda}$ or $\boldsymbol{\Lambda}$, just as it is done in the IQA

approach. Without loss of generality, we present an atomic partition of the TPS using the QTAIM. This can be made coarser (scaling it up to the functional group or molecular level) or finer (to the level of atomic core, bond and lone pair domains) at will. Another important point regards origin dependency, which may bring trouble in the partitioning if direct use of Eq. 6–1 is made. This difficulty is eluded by using the manifestly origin independent Eq. 6–4.

Let us start with an exhaustive partition of the physical space $R^3 = \bigcup A$ into atomic regions. A rigorous, physically sound possibility is provided by the QTAIM. Each of these regions or domains harbors a nucleus, at position $\mathbf{R}_A$. Given the one-to-one correspondence between domains and nuclei, we will label them interchangeably. Then,

$$\begin{aligned} \boldsymbol{\Lambda} &= \sum_{A \geqslant B} \boldsymbol{\Lambda}^{AB}, \qquad (6\text{–}7) \\ \boldsymbol{\Lambda}^{AA} &= \frac{1}{2}\int_A d\mathbf{r}_1 \int_A d\mathbf{r}_2 \, (\mathbf{r}_{12} \otimes \mathbf{r}_{12})\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2), \\ \boldsymbol{\Lambda}^{AB} &= \int_A d\mathbf{r}_1 \int_B d\mathbf{r}_2 \, (\mathbf{r}_{12} \otimes \mathbf{r}_{12})\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2). \end{aligned}$$

Notice that the above expressions provide a chemical partition of the TPS (or the LT if we divide by N).

The intra-atomic $\boldsymbol{\Lambda}^{AA}$ terms must tend to their free atomic values $\boldsymbol{\Lambda}_0^{AA}$ as the molecular system is pulled apart into atoms. Recalling that

$$\int d\mathbf{r}_1 d\mathbf{r}_2 \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = N,$$

it is well known that the localization index of region A,

$$N_{AA} = \int_A d\mathbf{r}_1 \int_A d\mathbf{r}_2 \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2),$$

defines the number of localized electrons in region A, so that $\boldsymbol{\Lambda}^{AA}$ measures the interelectron spread of these localized electrons in the atomic region, behaving grossly as $\boldsymbol{\Lambda}^{AA} \sim (N_{AA}/2)\langle r_{12}^2 \rangle_A$. The intra-atomic contributions to the TPS are thus additive and size extensive, and their sum is clearly seen through this partition to provide a term that scales linearly with the size of the system (or the number of electrons) as we approach the thermodynamic limit. An important corollary is that the root of any divergence in the LT will not be found in these intra-atomic components (see below). As it happens with other intra-domain expectation values in QCT, $\Delta\boldsymbol{\Lambda}^{AA} = \boldsymbol{\Lambda}^{AA} - \boldsymbol{\Lambda}_0^{AA}$ reflects the local change in interelectron spread due to chemical bonding and, except in very specific cases, like those in which a large charge transfer occurs, we expect these $\Delta\boldsymbol{\Lambda}^{AA}$'s to be small.

The interatomic $\boldsymbol{\Lambda}^{AB}$ contributions are much more interesting, for they directly measure the change in the interelectron spread due to the delocalization associated to the formation (or breaking) of a particular bond. From chemical

intuition, two spatially separated non-bonded atoms will display a vanishing $\boldsymbol{\Lambda}^{AB}$ value. Most, if not all, of the interesting behavior of $\boldsymbol{\Lambda}$ is then to be found in these terms. Notice that when two separated entities interact the Λ^{AA} components are non-vanishing at full-separation, changing upon interaction, but that the Λ^{AB} value is zero at infinite separation, sensing directly the interaction process.

6.3.3 How does Λ^{AB} decay with distance?

Taking into account that A and B are two non-overlapping regions of space, the behavior of $\boldsymbol{\Lambda}^{AB}$ as the two centers move away from each other is determined by the decay rate of ρ_{xc} with interelectron distance and the $\mathbf{R}_{AB} = \mathbf{R}_A - \mathbf{R}_B$ distance itself. Provided that the two electron coordinates satisfy $\mathbf{r}_1 \in A$ and $\mathbf{r}_2 \in B$, we may refer them to their local nuclear reference frames, respectively: $\mathbf{r}_1 = \mathbf{R}_A + \mathbf{u}_1$, $\mathbf{r}_2 = \mathbf{R}_B + \mathbf{u}_2$. With the above, the dependency of $\boldsymbol{\Lambda}^{AB}$ on the internuclear distance is explicitly separated. Let us define $\mathbf{u}_{12} = \mathbf{u}_1 - \mathbf{u}_2$, and the local integrals

$$\begin{aligned} \mathbf{I} &= \int_A d\mathbf{u}_1 \int_B d\mathbf{u}_2 \; (\mathbf{u}_{12} \otimes \mathbf{u}_{12}) \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2), \\ \mathbf{J} &= \int_A d\mathbf{u}_1 \int_B d\mathbf{u}_2 \; \mathbf{u}_{12} \; \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2), \end{aligned} \tag{6–8}$$

that may also be written in terms of spatial moments of the domain averaged Fermi holes introduced by R. Ponec [152], which have been successfully used in the last years to reveal many interesting effects in chemical bonding [53]. With these, we may write

$$\boldsymbol{\Lambda}^{AB} = \mathbf{I} + \mathbf{R}_{AB} \otimes \mathbf{J} + \mathbf{J} \otimes \mathbf{R}_{AB} + \frac{1}{2}(\mathbf{R}_{AB} \otimes \mathbf{R}_{AB})\delta^{AB}. \tag{6–9}$$

The first term contains only local distances, roughly decaying as δ^{AB} itself, and out of the three remaining terms, the one leading the long-range behavior is the third. Thus, at large interatomic distances $\boldsymbol{\Lambda}^{AB} \sim (\mathbf{R}_{AB} \otimes \mathbf{R}_{AB})\delta^{AB}/2$, and the parallel component of $\boldsymbol{\Lambda}$ along the bond direction will scale as

$$\Lambda_{\parallel}^{AB} \sim \frac{1}{2} R_{AB}^2 \delta^{AB}. \tag{6–10}$$

This last important relation provides a new bridge between TCB descriptors in the ground state and the Kohn-Resta theory of the insulating state.

6.3.4 The chemical bonding origin of the convergence/divergence of λ

We can now turn to the convergence/divergence of $\boldsymbol{\lambda}$ in the thermodynamic limit. An explicit effective one-center expansion of Λ may be immediately written from our previous partition as

$$\boldsymbol{\Lambda} = \sum_A \boldsymbol{\Lambda}^A,\ \boldsymbol{\Lambda}^A = \boldsymbol{\Lambda}^{AA} + \frac{1}{2}\sum_{B\neq A} \boldsymbol{\Lambda}^{AB}. \tag{6–11}$$

Notice that the atomic additivity of $\boldsymbol{\Lambda}^A$ allows us to write

$$\boldsymbol{\lambda} = \frac{1}{N}\sum_A^{N_{at}} \boldsymbol{\Lambda}^A = \frac{N_{at}}{N}\langle \boldsymbol{\Lambda}^A\rangle = \frac{\langle \boldsymbol{\Lambda}^A\rangle}{n}, \tag{6–12}$$

where $\langle \boldsymbol{\Lambda}^A\rangle$ is the average of $\boldsymbol{\Lambda}^A$ over all the atoms comprising our system, and n is the average number of electrons per atom, The divergence of $\boldsymbol{\lambda}$ in the thermodynamic limit is equivalent to that of the average atomic-additive $\langle \boldsymbol{\Lambda}^A\rangle$. Remarkably, the divergence of the LT can thus be nailed down to an atomic property.

Further analysis opens new avenues in understanding the onset of conductivity from a chemical perspective. Several paths may lead to a divergent $\langle \boldsymbol{\Lambda}^A\rangle$. For instance, all of the $\boldsymbol{\Lambda}^A$ terms may diverge themselves, or only one or a few. This analysis will identify *essential and spectator* atoms or functional groups in complex conducting systems. Essential groups for conductivity will be those for which $\boldsymbol{\Lambda}^A$ diverges, while spectator groups will be characterized by convergent $\boldsymbol{\Lambda}^A$. We think that this classification scheme can help identify replaceable groups that will not change the basic conductivity properties of a system while tuning their fine conductive properties.

For each divergent $\boldsymbol{\Lambda}^A$, our previous comments show that it will be the interatomic sum, $\sum_{B\neq A} \boldsymbol{\Lambda}^{AB}$, not the intra-atomic $\boldsymbol{\Lambda}^{AA}$, that will add to an infinite result. It is the interplay between the dimensionality of the system and the decay rate of δ^{AB}, that determines convergence. This binds the behavior of $\boldsymbol{\lambda}$ to the decay rate of DIs, already explored [61, 62].

To keep our discussion as simple as possible, we will now continue our reasoning in one-dimensional systems, where these ideas are most easily aprehended. In 1D, whenever δ^{AB} decreases faster than $\delta^{AB} \simeq R_{AB}^{-d}$, with $d = 2$, then the $\sum_{B\neq A} \boldsymbol{\Lambda}^{AB}$ term will converge, and the contrary will make it diverge. Similarly, the limiting d exponent is 3, 4 for 2- and 3-dimensional conductivity to occur, respectively. These results perfectly match the findings relating the decay rate of the non-diagonal elements of the first order density in tight binding models of metals, as shown by Taraskin [183, 184]. On the contrary, exponentially decaying interatomic delocalization indices δ^{AB} will always lead to insulating behavior, i.e. to convergent $\boldsymbol{\lambda}$ values. We would like to stress that the transition from exponential to power-law δ^{AB} decay rates has already been found to occur in computational studies of model systems [62].

The following expressions summarize the core of our findings in the thermodynamic limit:

- If δ^{AB} decays exponentially with R^{AB}, $\boldsymbol{\lambda}$ converges.
- In a d-dimensional system, $\boldsymbol{\lambda}$ converges/diverges if δ^{AB} decays faster/slower than $R_{AB}^{-(d+1)}$.

We propose that low-dimensional, as well as bulk conductivity can be spotted by examining the behavior of $\boldsymbol{\Lambda}^{AB}$ along the appropriate directions, planes, or 3D regions, respectively.

6.4 EXEMPLIFYING THE USEFULNESS OF THE Λ PARTITION

We are now in a position to show the new insights at work in a few toy systems. We will first discuss two simple dissociation processes, $H_2 \rightarrow H + H$ and $H_2O \rightarrow OH + H$, where we will see how, and why, the final localization of electrons in the products leads to a convergent $\boldsymbol{\lambda}$, i.e. to insulating-like behavior. This will also make the essential role of electron correlation in correctly capturing the physics of the system. Finally, the early stages in the birth of a divergence, thus the switch towards metallic-like behavior, will be succintcly analyzed in a linear chain of equally spaced hydrogen atoms. The electronic structure calculations have been performed with the `GAMESS` package [166], and the TPSs have been obtained for QTAIM atomic partitions through our `PROMOLDEN` [107], code, which is able to handle quite a number of correlated and non-correlated wave functions and several QCT partitions, not only the one provided by the QTAIM. Details of the implementation of the TPS in `PROMOLDEN` can be found in the ESI.

6.4.1 The dissociation of H_2

First we discuss the H_2 molecule ($A = H$, $B = H'$), a paradigm of covalent interactions. We have computed $\boldsymbol{\Lambda}$ at the Hartree-Fock (HF) and the configuration active space (CASSCF) levels with the aug-cc-pVTZ basis set along its dissociation coordinate. The results are contained in Fig. 35. Notice that the $\boldsymbol{\Lambda}$ tensor is diagonal in any reference frame in which the internuclear distance coincides with one of the coordinate axes, and that rotational invariance equalizes the other two orthogonal eigenvalues of $\boldsymbol{\Lambda}$. We will call these two different components of $\boldsymbol{\Lambda}$ $\Lambda_{\parallel}$ and $\Lambda_{\perp}$, respectively. As already put forward by Resta [162] and Leininger *et. al.* [29], mean-field and correlated descriptions of the dissociation process differ essentially. Interestingly, this qualitative differences are also observed when the delocalization index is examined by itself [108].

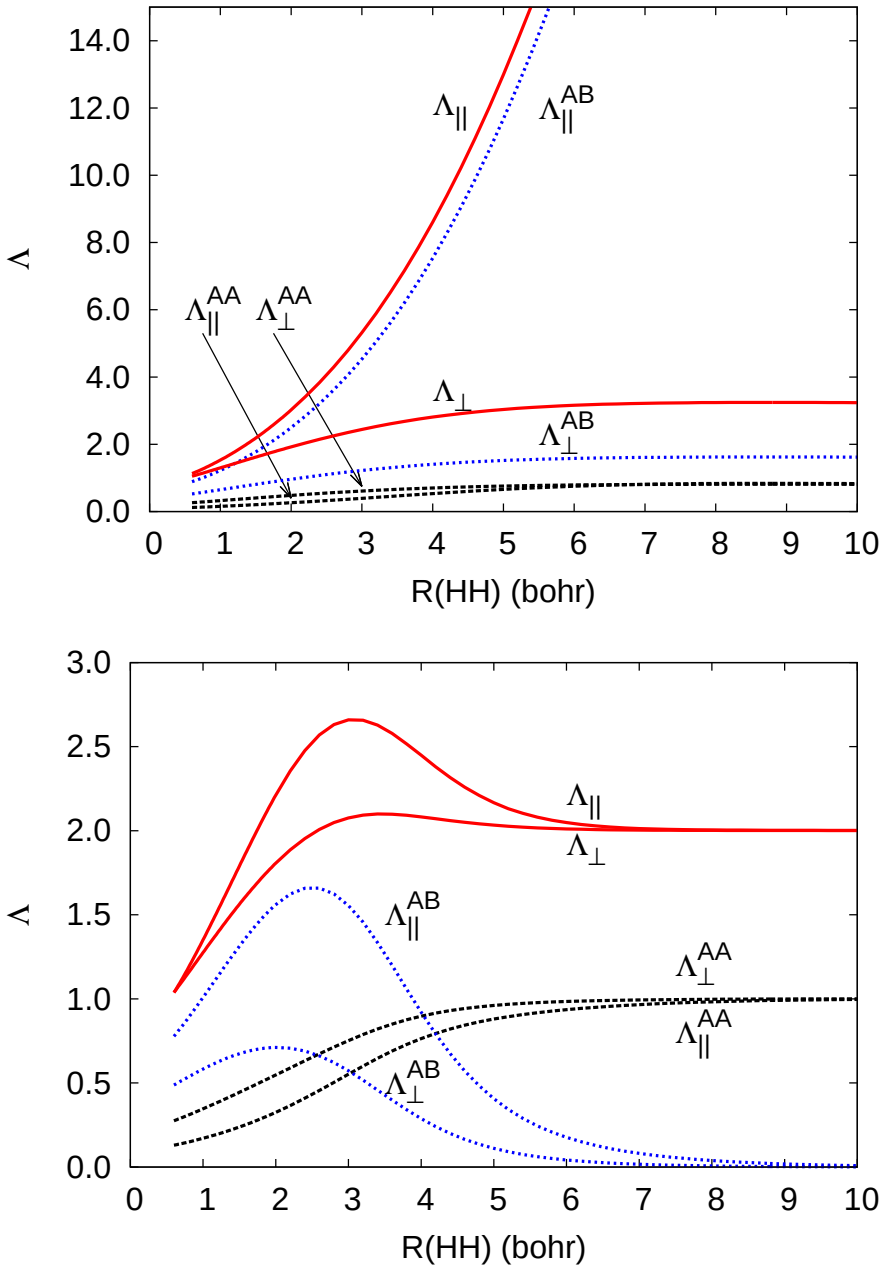


Figure 35: Total, intra-atomic, and interatomic components of Λ in the H_2 molecule at the HF (top) and CASSCF//aug-cc-pVTZ (bottom) levels along the internuclear dissociation coordinate. The parallel and perpendicular labels correspond to the internuclear and orthogonal directions, respectively. All data in au.

Failure to consider electron correlation leads to a parabolic divergence of $\Lambda_{\parallel}$ as the internuclear distance increases. Its origin cannot be grasped by solely examining the full tensor, but its partitioning shows that, as expected, it is the interatomic component, $\Lambda_{\parallel}^{AB}$, that diverges. Through the eyes of our findings, the parabolic behavior is due to an artificial non-vanishing δ^{AB} at infinite separation (the HF dissociation error). It may be instructive to recall that at the Hartree-Fock level we may write for closed-shell systems

$$\delta^{AB} = 4 \sum_{ij} S_{ij}^{A} S_{ij}^{B}, \tag{6–13}$$

where the sum runs over all pairs ij of occupied orbitals and S_{ij}^{A} is an atomic overlap integral,

$$S_{ij}^{A} = \int_{A} d\mathbf{r}\, \phi_i(\mathbf{r}) \phi_j(\mathbf{r}). \tag{6–14}$$

As in H_2 we have only one occupied HF orbital fulfilling $S_{11}^{A} = S_{11}^{B} = 1/2$ by symmetry considerations, the bond order $\delta^{AB} = 1$ at any R_{HH} in this model. This can be also interpreted as the result of the two opposite spin electrons being statistically independent if no Coulomb correlation is added. The wrong constant δ^{AB} leads to infinite-range delocalization, with an overall probability of finding the two electrons in any one of the H atoms (the so-called ionic weight in quantum chemical approaches) equal to 1/2 [52]. Eq. 6–10 does the rest.

Proper inclusion of Coulomb correlation makes δ^{AB} decrease exponentially at large distances [111], so that the bond breaks appropriately. Thus, in the correct correlated description, the intra-atomic $\boldsymbol{\Lambda}^{AA}$ components start at low values close to equilibrium, increasing to the free atom limit. From Eq. 6–1, it is clear that in this limit $\boldsymbol{\Lambda}^{AA}$ is also diagonal and that each of its three components is equal to $\langle \phi | r^2 | \phi \rangle$, where ϕ is the hydrogenic atomic orbital in the state of interest. For the 1s ground state, $\langle r^2 \rangle = 3$ au. so that $\langle z^2 \rangle = 1$, which is also the appropriate $\boldsymbol{\Lambda}^{AA}$ limit at dissociation.

Many other features of the behavior of the intra-atomic components are easy to rationalize. For instance, the lower value of $\boldsymbol{\Lambda}^{AA}$ at small interatomic distances is understood straightforwardly, for the number of localized electrons in each atom is in these conditions about 0.5, *vide supra*. It is also interesting to notice that the intra-atomic $\Lambda_{\parallel}^{AA}$ value is smaller than its $\Lambda_{\perp}^{AA}$ counterpart, this reflecting the compression of the atomic density along the internuclear axis as we approach the two atoms from infinity. As also expected, it is the interatomic $\boldsymbol{\Lambda}^{AB}$ component that accounts for the sharp maximum in the total $\boldsymbol{\Lambda}$ tensor. This maximum has been interpreted [29] as a signature of bond breaking. Under our present formalism, it is a simple consequence of the shift from a power-law to an exponential decay in ρ_{xc} or δ^{AB}, i.e. from a quasi-independent electron pair being stretched (as in the HF case) to the strongly correlated, localized dissociation limit. This transition, scaled by $R_{AB}^2/2$, gives rise to the

maximum, that is found very close to the internuclear distance at which the inflection point of δ^{AB} has been repeatedly described [62, 65].

6.4.2 The power of partitioning Λ: $H_2O \rightarrow OH + H$

We can now show how the global behavior of $\boldsymbol{\Lambda}$, that contains the total response of the system to a chemical process, may be split up into chemically meaningful terms. To that end we have chosen the $H_2O \rightarrow OH + H$ dissociation, with the OH_1 distance taken as an intrinsic reaction coordinate. Fig. 36 shows the evolution of $\Lambda_{\parallel}^{AB}$ along R(OH1) at the complete active space CASSCF[8,8]//aug-cc-pVTZ level. This exemplifies the power of partitioning Λ. As the OH1 distance is stretched, we see how there is a simple jump in $\Lambda_{\parallel}^{OH_2}$, while it is $\Lambda_{\parallel}^{OH_1}$ that behaves much as in the H_2 case. The step from lower to higher $\Lambda_{\parallel}^{OH_2}$ can be understood by taking into account that in the final OH radical the number of delocalized electrons between the O and the H2 atoms has increased. In other words, since the OH2 bond order increases as the H1 atom dissociates, $\Delta\delta^{OH_2} > 0$, so does $\Lambda_{\parallel}^{OH_2}$. Only a partitioning of the TPS, like the one devised here, will be able to isolate the main actors in complex scenarios. With our tools, this seems to be at hand, and the strong link between the essential interatomic $\boldsymbol{\Lambda}^{AB}$ terms and the DIs is unveiled.

6.4.3 Recognizing the onset of conductivity: The H_{10} chain

Our next example will be a linear chain of 10 equally spaced H atoms computed at the HF and full valence CASSCF levels with the 6-311G* basis set. At the inter-hydrogen distance selected, $R = 3.5$ bohr, the HF model is starting to fail, but it still provides a reasonable description of the electron system. We examine how the interlectron spread propagates along a quasi-1D system, and our goal is put on Eq. 6–11. Fig 37 shows how $\Lambda_{\parallel}^{AB}$ changes for all pairs in which one of the atoms is fixed to be an end H. The first interesting point is that in the mean-field HF approximation, $\Lambda_{\parallel}^{AB}$ decays with distance in a well-developed slow oscillatory pattern. We have found a similar behavior examining DIs [61, 62].

Notice that, from the chemical point of view, these oscillations signal a clear bond order alternation or mesomerism, precursor of a Peierls distortion (or H_2 dimerization). This has also been repeatedly described in previous literature: [65] delocalization indices in geometrically constrained systems inform about the expected distortions when the constraints are released. As we find here, the localization tensor yields similar sensible chemical information.

A power law fit of δ^{AB} to R_{AB}^{-d} gives $d \approx 2.5$, close to the tight binding value ($d = 2.0$). The sum in Eq. 6–11 achieves a very large value. The oscillatory pattern in $\Lambda_{\parallel}^{AB}$, as in the case of the DI, is a clear indicator of conducting-like behavior. A second point regards the very quick saturation of $\boldsymbol{\Lambda}^{AB}$ to the

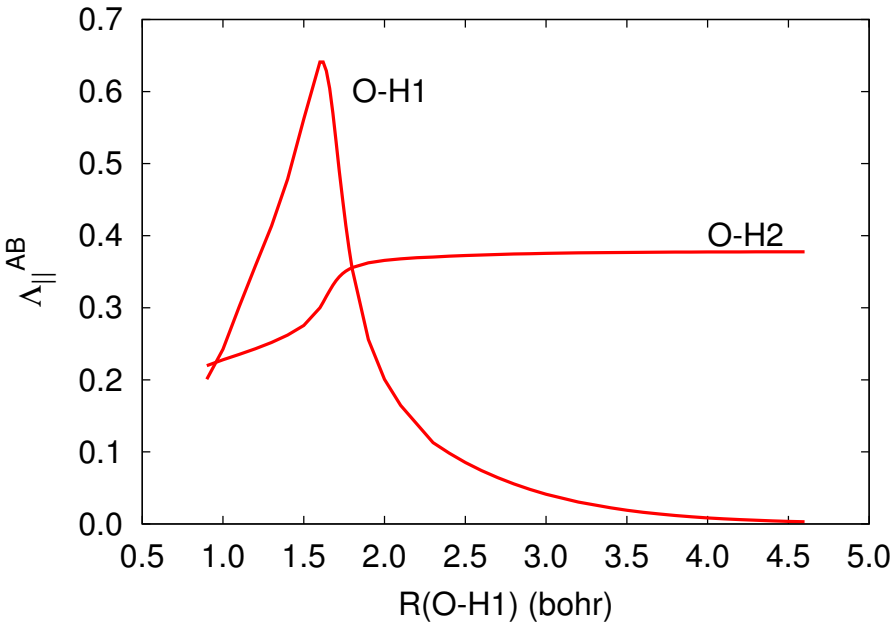


Figure 36: Interatomic components of $\Lambda_{\parallel}$ in the H_2O molecule at the CASSCF[8,8]//aug-cc-pVTZ level along the internuclear O-H_1 dissociation coordinate. All data in au.

$R_{AB}^2 \delta^{AB}/2$ leading term in the long range. Fig. 37 shows that our previous theoretical insights are fully realized from actual computations. It is the decay rate of DIs (i.e. the inter-center electron delocalization) that determines conductivity in the thermodynamic limit. Inclusion of electron correlation does not make the oscillations disappear at this interatomic distance, but reveals how the electrons are now much more localized, with a considerably smaller spread. A similar fit now gives $d \approx 4.1$, well above the metallic limit. Previously it was shown that the oscillations disappear when we enter the dissociating, localized regime, and that the DI decays exponentially in that case [62].

6.4.4 Insulator-like and conducting-like chains

We will finish our discussion by considering two real life one-dimensional linear chains of equidistant atoms: $(LiH)_{15}$ and Li_{10}, with nearest neighbor distances set to 3.0 and 5.818 bohr, respectively. We have used a HF/6-311G* level that provides a simple, yet reasonable description of both systems. Fig. 38 shows relevant values for $\Lambda_{\parallel}^{AB}$, that provide clear grounds for comparison: the small values and the very quick decay of the interatomic $\Lambda_{\parallel}^{AB}$ values in an

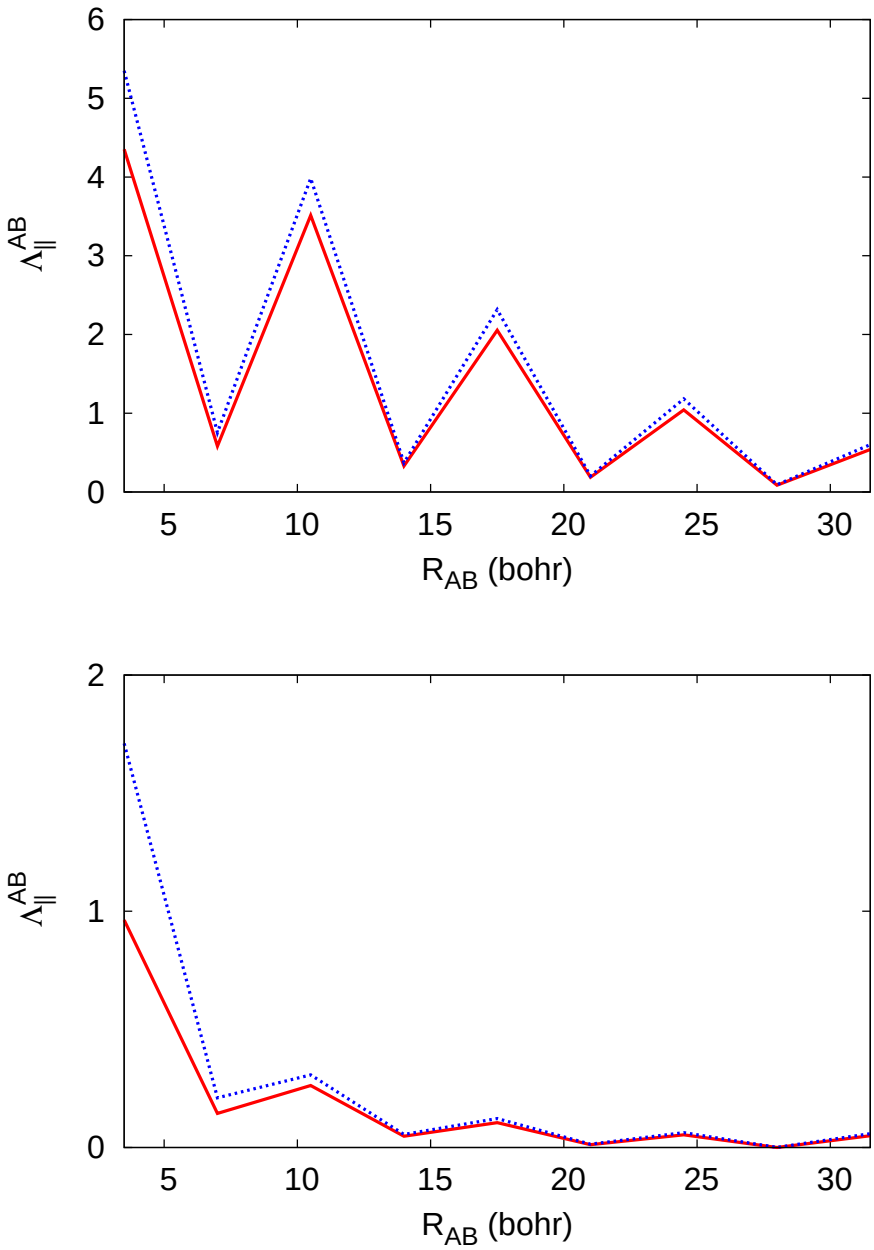


Figure 37: $\Lambda_{\parallel}^{AB}$ along the internuclear direction in a H_{10} linear chain of equally space atoms at the HF//6-311G* (top) and CASSCF//6-311G* (bottom) levels. A labels one of the end H atoms, and B runs over all the others. Nearest neighbohrs are 3.5 bohr apart. In each plot, the dashed curve corresponds to $R_{AB}^2 \delta^{AB}/2$. All data in au.

insulator like lithium hydride, and their much larger magnitude, slow decay, and oscillatory behavior in the metallic-like chain.

6.5 CONCLUSIONS

Summarizing, we have shown in this chapter that a remarkable bridge exists between the Kohn-Resta theory of the insulating state, through the localization tensor (or its total position spread tensor version in molecular systems), and well known indicators used in the modern theory of chemical bonding as bond orders. This has been achieved by partitioning the localization tensor in intra- and interatomic components. An orbital invariant way to do so starts by writing the LT in terms of reduced densities and then partitioning the space into atomic regions according to quantum chemical topology. Convergence or divergence of the LT in the thermodynamic limit , associated to insulating or conducting electrical properties, depends exclusively on the decay rate of its interatomic components. The latter are dominated by the chemical delocalization index, a modern form of bond order. The chemistry of ground states and the physics of conductivity become intertwined in this way. We expect this new link to be useful in the search and design of low dimension conductors or insulators, for the total LT can be written as a sum of atomic (or functional group) components. Each atom or functional group in a system may thus be classified as essential, if its contribution to the LT diverges, or an spectator, if it converges, as electrical conductivity is regarded. We expect that this categorization can be used advantageously in the rational design of new materials.

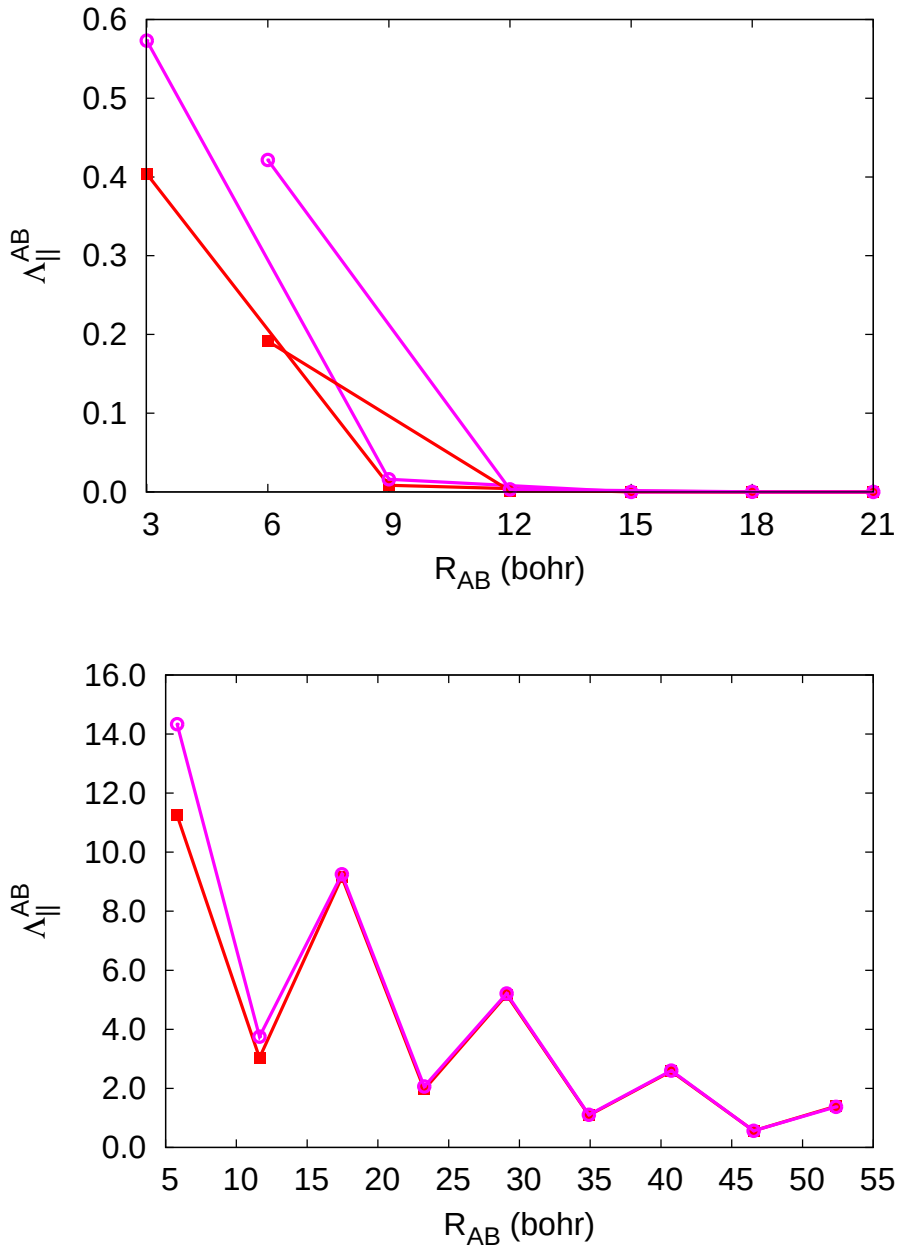
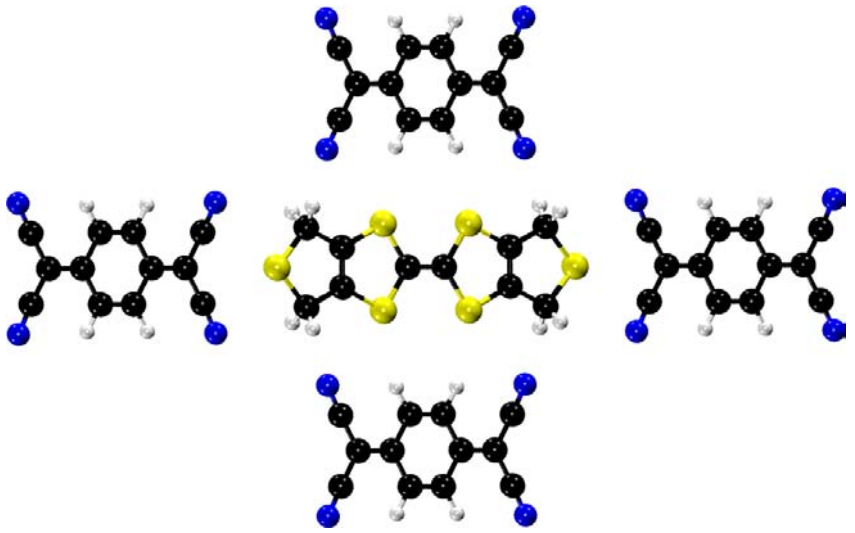


Figure 38: HF//6-311G* $\Lambda_{\parallel}^{AB}$ (bold squares) and $R_{AB}^2 \delta^{AB}/2$ (circles) along the internuclear direction. Top: linear $(LiH)_{15}$ chain of equally spaced nearest-neighbors atoms. Values starting at $R_{AB} = 3.0/6.0$ bohr correspond to the AB=LiH/HH pairs, respectively. Bottom: linear Li_{10} chain of equidistant Li atoms separated by 5.818 bohr. The AB pairs are those formed by one end Li atom and all of its neighbors.

7 TOPOLOGY OF THE ELECTROSTATIC POTENTIAL IN SOLIDS

CONTENTS



BTDMTTF–TCNQ is a modification of the historically first organic crystal conductor found.

7.1 INTRODUCTION

The use of the molecular electrostatic potential (MEP) as a medium for predicting attractive and repulsive interactions between molecules has become part of mainstream chemistry, being very common to find MEP maps depicted in organic chemistry textbooks [**?**]. The rationale behind those characteristic images that project the MEP on low density isosurfaces is that one can identify nucleophilic and electrophilic regions — useful when predicting regioselectivity. Owing to the unique property of the potential to emphasize the effect of nuclear charges in non-covalent interactions. Aside, this has its limitations, it is possible only in those cases where polarization and charge transfer can be neglected.

During the last decades plenty of works highlighted applications in molecular reactivity and host-guess recognition [**? ?**]. However, very extensive use of the MEP has been made on a more qualitative level. On this regard the topological analysis [**?**] provides a description of the electrostatic characteristics from a reduced set of points in space without relying on the choice of an arbitrary surface. The electrostatic potential is a one-particle observable, with physical meaning, that is compliant with topological approaches. The characterization of the MEP in molecular systems [59, 60, 97], headed by Gadre and coworkers, gave rise to the first quantitative indices. Electronegativity and covalent radii, two deeply rooted concepts in chemistry, can be recovered [151] by a formulation in terms of the properties at critical points. Also, the domain of nucleophilic and electrophilic influence zones was established with the partition of space from the electrostatic potential [114]. Nevertheless, a full characterization of the topology of the electrostatic potential is still far from being complete.

Besides, extended systems have been left apart in the development of the electrostatic potential analysis. The inner crystal electrostatic potential (ESP) could be particularly useful for designing new materials. It would be convenient to connect the properties of the electrostatic potential with an elucidation of structural stability. Electron diffraction experiments provide the ESP of strongly bonded solids such as ionic [72] or covalent solids but molecular systems, where the long range interactions stabilize the structure, cannot be examined with this technique. For that purpose X-ray diffraction experiments provide the density and from it we can obtain the ESP for a moiety that resembles the properties of the solid. Up to now, few molecular crystals underwent a topological analysis of its experimental electrostatic potential [28, 114].

High quality single crystals are needed to characterize electronic properties in those systems. Crystal defects and undesired variations in the quality of the structure factor amplitudes introduced by other experimental effects can be overcome with the support of ab initio data. Of course, before such theoretical predictions can be trusted, their accuracy in reproducing experimentally determined data must be ascertained. The electrostatic potential of BTDMTTF-TCNQ has been characterized experimentally, and is a good starting point in this case. We considered appropriate the use of the electrostatic potential to

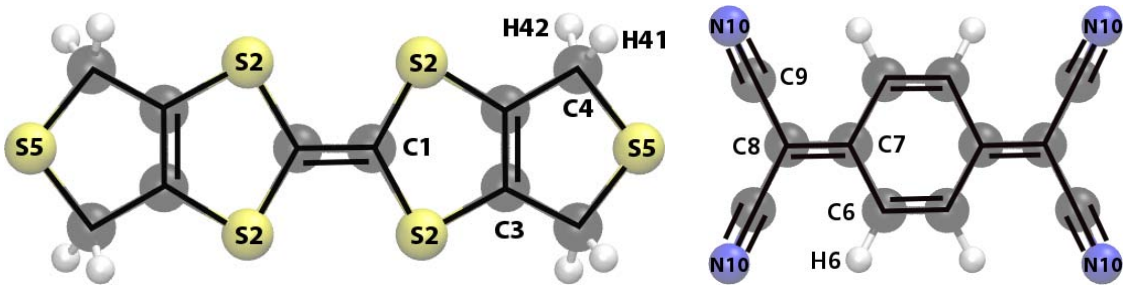


Figure 39: BTDMTTF and TCNQ chemical schemes. The charge transfer magnitude is unknown. It has been previously approximated from multipolar parameters.

study an organic conductor in crystal phase with the aim of extending the previous work by Mata et al. BTDMTTF-TCNQ follows the lead of other milestone organic conductors like TTF-TCNQ and $(TMTSF)_2PF_6$ for being the first organic metal and superconductor respectively. BTDMTTF-TCNQ has the remarkable characteristic of being an almost pure 1D metal down to 26 K. Organic conductors offer a series of advantages over the more traditional pure metals that can be summarized into: easier manufacturing process; lower production cost; more environmental friendly; and their mechanical flexibility can have useful applications.

Our objective is to further investigate the topology of the ESP in the solid state system BTDMTTF-TCNQ to understand its physical properties. BTDMTTF-TCNQ belongs to a class of organic charge-transfer systems (that can be realized as high conducting materials) for which most properties are determined solely by the degree of charge transferred between the donor and acceptor molecules. Controlling this variable would be a desirable capacity. The idea is to recognize the interactions responsible of the resulting charge transfer to design new materials with desired properties. Above all, non-covalent interactions provide the most important contribution to the binding energy in molecular solids. The additional Coulomb and dipole contribution to the binding energy in this type of compounds is better examined with the topology of the electrostatic potential.

First of all we will present what is known about its structure. The most stable crystal phase at 130 K has been characterized with single crystal X-ray diffraction. The structure belongs to a monoclinic space group type: $C_{2/m}$ with molecules arranged as columns along the c axis. Each column has molecules of only one type and is surrounded by other four of the opposite molecule (figure 40). They are not stacked at the same c elevation, they are displaced one with respect to another, and there is a small angle ($\approx 15^o$) between them. The conductivity would take place after BTDMTTF–TCNQ has undergone an excitation to a state where BTDMTTF and TCNQ molecules have, formally, a positive and a negative charge, respectively. (see figure 39). Holes move along columns of donor molecules and electrons along acceptor columns (see figure **??**).

bis(thiodimethylene)–tetrathiafulvalene (BTDMTTF)

7,7,8,8–tetracyanoquinodimethane (TCNQ)

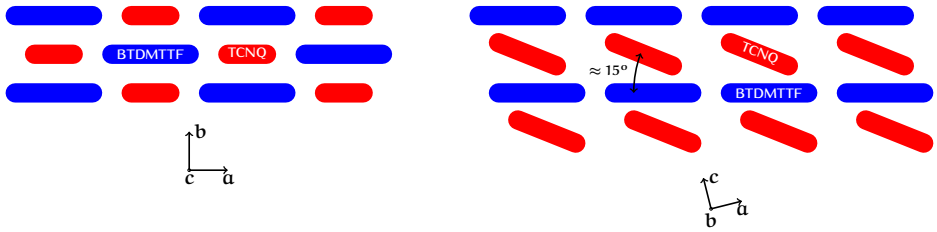


Figure 40: Schematic structure of BTDMTTF–TCNQ.

A previous work on the system obtained net atomic charges from kappa and multipolar model parameters. The results thereof [46] reported an estimate of the charge transferred upon excitation (q_{CT}) around 0.7 e. However, that value should be taken with care since those are models more tied to fit the experimental density than to define an atom in a molecule. Instead an integration over topological basins would provide atomic charges with a quantum mechanical grounding.

To design new materials with enhanced conductivity we need also to know why there is charge transference. There are several questions that should be answered: what electrostatic interactions are present? If there is any interaction that plays a greater role than the others, how are they involved in the charge transfer process? How one may hope to modify them to improve material's performance? Or how they stabilize the whole crystal structure? Again, a topological analysis can shed light on this matter. Aside from the well studied density, the electrostatic potential is the best suited to examine the increased Coulomb and dipolar interactions. There is no evidence in the literature to determine what interactions are present, how are they involved in the charge transfer process or how they stabilize the whole crystal structure. In a preliminary article that inspected the electron density map it remarked that most of the charge transfered is concentrated in nitrogen atoms. And that the external sulfur (S_5) to C≡N triple bond interactions as the main charge transference pathways. Again, a topological analysis can shed light on this matter.

We will take advantage of electron density, electron density Laplacian, and electrostatic potential topology.

7.2 THEORY

As a novelty we will contrast the basins of the density and the electrostatic potential to gain more insight about the charge transfer. In this respect, ESP basins reflect the deformation, and expansion/compression, inflicted by electrostatic interactions with the environment to the charge distribution of the reference, neutral, spherically symmetric, isolated atoms that originally fulfill the whole space. They do not display charge transfer effects, only the defor-

mation. Although this perspective introduces an arbitrary reference state (it is not known if it should be the ground state or not) it might be a useful point of view to get qualitative information of the behavior of the atoms. As opposed to the electron density basins, ESP basins contain the same charge as the isolated atom. When we switch from the ESP basin to the density basin we travel between two images of an atom include the effect of charge transfer.

Assuming that the electron density and electrostatic potential zero-flux surfaces are related by a scaling function, that is, basins of same shape but of differing size, the space enclosed by both surfaces allows us to localize atom averaged partial charges. If $\Lambda_A \subset \Omega_A$ (anion) is satisfied then:

$$\int_{\Omega_A \setminus \Lambda_A} \rho(\mathbf{r}) = \delta_A^- \tag{7–1}$$

Otherwise if $\Omega_A \subset \Lambda_A$ (cation) then:

$$\int_{\Lambda_A \setminus \Omega_A} \rho(\mathbf{r}) = -\delta_A^+ \tag{7–2}$$

The potential due to a charge ρ outside of the charge distribution is much simpler to evaluate if we consider the contribution of those basins. The convergence issue is ameliorated.

7.3 METHODOLOGY

DFT calculations were performed with a plane-wave basis set using pseudopotentials that adopt the all-electron projector augmented wave (PAW) method [22]. We used the hybrid HSE06 functional [94] as implemented in the Vienna *ab initio* package (VASP) [93] to account for corrections of the van der Waals forces. The plane wave cutoff was set to 520 eV and the reciprocal space mesh was 4×4×6 to achieve a good compromise between available resources and high accuracy. Experimental cell parameters and atomic coordinates from previous experimental charge density refinements [47] were used as input for the VASP calculation and were later optimized (a =).

The experimental charge density was acquired applying a multipolar pseudoatom model [77] with the MOPRO suite [74, 86**?**]. The labeling of non equivalent atoms is presented in figure **??**. Starting from the density, one can proceed to evaluate the electrostatic potential. The accuracy of the resulting potential depends on the level of approximation used for the charge density.

`MoProViewer` and `critic2` [135, 138] allowed us to analyze critical points and integrate electron density basins. For the experimental densities, since the multipolar parameters are fitted against the density our approach consists in selecting a moiety, contributions from other atoms to the density are neglected. We are analyzing atoms isolated but the multipolar parameters are those of the crystal. Appendix **??** explains the steps followed for the construction of

the pseudo-isolated molecules. However, some density and electrostatic potential basins of isolated molecules are open. The choice of a cluster of an ion embedded inside its nearest neighbors solves the problem.

We identified more (3,-1) critical points in the optimized geometry than in the experimental geometry that are equivalent to those resulting from the Hansen-Coppens density. So we choose the former geometry from now on to realize comparative studies.

Basin integrations were performed using both Yu-Trinckle [191] and Henkelman [78, 182**?**] algorithms, releasing quite the same values, being the Yu-Trinckle integration a little more accurate. The integration of electron density basins was quite reasonable, in close agreement with experimental charges.

Taking into account that the shape of the electrostatic potential of a molecule should equal roughly that of a solid in the internuclear region it was found that the electrostatic potential that VASP exports does not satisfy this condition for unknown reasons. On the other hand QUANTUM ESPRESSO [68] outputs a potential that obeys the stated condition. There is no HSE06 functional available in quantum espresso but the electrostatic potential is obtained from the density and this one is not very much affected by the functional of choice. If we consider HSE06 as a perturbation of PBE, according to Wigner's rule, only second order corrections to the energy affect to first order to the wavefunction and thus the density. We studied the electrostatic potential calculated with quantum espresso that contains corrections to van der Waals forces (PBE [149]+XDM [19]).

The quality of $\phi(\mathbf{r})$ basin integrations has been tested by checking the value of the integrated density within the ϕ-basins, which must be 0. For each atom the basin charge was smaller than 8×10^{-2} e, indicating that only modest errors are present in the integrated properties.

We choose atomic units except when otherwise stated.

The way to obtain the electrostatic potential is either from X-ray diffraction experiments or from ab initio calculations. We cannot obtain directly the electrostatic potential as a Fourier transform of the structure factors because the ESP has a singularity when the reciprocal vectors are 0. We can't obtain the zero potential. So, we no longer talk about the electrostatic potential of the whole solid, instead we take a moiety of the crystal and from the density of that moiety we obtain the electrostatic potential. We analyze so-called pseudo-isolated molecules. Nevertheless the parameters are those of the crystal. It is shifted by a constant but it is not important because we are interested in CPs, in relative values.

We will analyze several pseudo-isolated molecules: a) each molecule pseudo-isolated b) several dimers to find out what interactions stabilize the structure c) and finally a cluster of each molecule embedded inside its nearest neighbors to achieve closed basins.

7.4 RESULTS AND DISCUSSION

7.4.1 Intramolecular interactions

When electrons are promoted from donor molecules to acceptor molecules, the cations are stabilized by two resonant structures, one of them is proposed to be more stable due to $4n + 2\pi$ electrons located in its cycle. That structure has internal sulfurs (S2) with sp^2 hybridization. Does it happen the same in this structure? The ratio of electron density field curvatures at BCPs, ellipticity, is used as an indicator of the π character of those interactions.

Property values at (3,-1) critical points of BTDMTTF are collected in table 13. The low values of the density at the BCPs (S2–C1, S2–C3, S5–C4) emphasize the identity of sulfurs as elements isolated from the rest of the molecule. External sulfurs are more isolated than internal sulfurs. The Laplacian is negative indicating that it is a covalent interaction, but the small values indicate that they are close to a non covalent interaction. The electrostatic interaction component is also smaller than other interactions present in the molecule.

The only one interaction with a significant π character is C3-C3'. We see that the original double bond C1-C1 of the neutral BTDMTTF has disappeared.

Nor the internal neither the external sulfur present any π bonding character. It suggests that both have sp^3 hybridization.

The analogy of BCPs and EBCPs is broken by an interaction between internal sulfurs of relative long range.

ESP (3,+3) critical points located in front of nitrogens denote lone pairs location. In the presence of neighbor molecules they disappear.

7.4.2 Intermolecular interactions

What chemical interactions are present and which one are playing the mayor role? Which is their nature and how one may hope to modify them to improve material's performance? Are these interactions localized or not?

There are two types of intermolecular interactions between columns: frontal and lateral. Frontal interactions involve the external sulfurs (S5) and lateral interactions involve the internal sulfurs (S2), figure 43. The LONGITUDINAL pseudo-isolated dimer is a BTDMTTF cation and its nearest frontal TCNQ anion. The LATITUDINAL pseudo-isolated dimer is a BTDMTTF cation and its nearest lateral TCNQ anion. External sulfurs are displaced out of the cation plane and they are nearly coplanar with the anion. Such displacement is originated by its interactions with TCNQ. It is reasonable to guess that frontal interactions are stronger based upon this observation. Otherwise, we need a more rigorous criteria to assert if any interaction is predominant.

External sulfurs interact with the carbonyl triple bond of the nearest anion, figure 44 (a) interactions, and internal sulfurs interact with the nearest nitrogen (c) and with hydrogens of anions that are above and below (d and e). The sigma

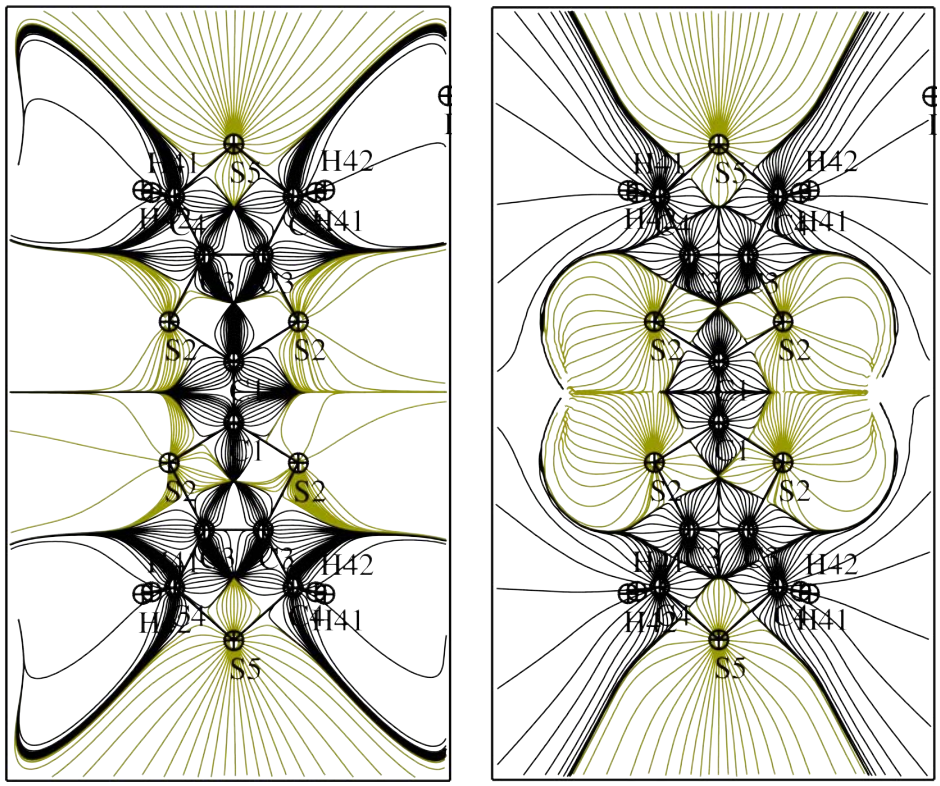


Figure 41: BTDMTTF: Electronic Density (ED) gradient lines (left) follow to the infinite indicating that the molecule is positively charged while some basins of the ESP (right) have a finite surface limit enclosing zero net charge. S5 atoms are not in the plane.

Table 13: BTDMTTF and TCNQ: Properties of the ED (top block), ρ, and ESP (bottom block), ϕ, scalar fields at $(3,-1)$ critical points between nuclei A and B. d_A and d_B are the distances from the critical point to the positions A and B taken as the length of the straight lines connecting A or B to the CP. Thus $d_{AB} = d_A + d_B$. $\lambda_i(e/Å^5)$ and $\gamma_i(e/Å^3)$ are the curvatures of the density and the electrostatic potential, respectively. As already mentioned in chapter 3 the ordering of curvatures is $\lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$; $\gamma_1 \leqslant \gamma_2 \leqslant \gamma_3$.

ρ: A—B	d_A(Å)	d_B(Å)	d_{AB}(Å)	$\rho(e/Å^3)$	$\nabla^2\rho$	λ_1	λ_2	λ_3	ε
S2—C1	0.968	0.845	1.813	1.275	−0.813	−6.484	−5.483	11.154	0.183
S2—C3	0.849	0.888	1.733	1.341	−2.705	−7.200	−6.240	10.735	0.154
C3—C4	0.728	0.746	1.473	1.765	−10.622	−11.779	−10.427	11.584	0.130
S5—C4	0.960	0.834	1.793	1.131	−0.320	−5.601	−5.205	10.486	0.076
C4—H41	0.794	0.228	1.022	1.667	−18.366	−18.458	−15.763	15.855	0.171
C4—H42	0.795	0.224	1.019	1.675	−16.427	−17.158	−15.669	16.400	0.095
C1—C1′	0.688	0.688	1.376	1.985	−11.046	−12.396	−10.27	11.62	0.207
C3—C3′	0.674	0.674	1.348	2.457	−21.660	−19.705	−13.424	11.469	0.468
N10—C9	0.735	0.424	1.159	3.682	−19.684	−30.801	−27.887	39.004	0.104
C8—C9	0.669	0.748	1.418	1.871	−9.597	−13.042	−10.725	14.170	0.216
C7—C8	0.682	0.718	1.400	1.814	−8.165	−11.442	−10.069	13.346	0.136
C7—C6	0.686	0.750	1.436	1.919	−11.944	−14.027	−10.976	13.059	0.278
C6—H6	1.076	0.037	1.100	1.783	−18.246	−18.757	−14.841	15.352	0.264
ϕ: A—B	d_A(Å)	d_B(Å)	d_{AB}(Å)	$\phi(e/Å)$	$\nabla^2\phi$	γ_1	γ_2	γ_3	
S2—C1	1.030	0.784	1.814	1.163	17.665	−5.192	−5.077	27.934	
S2—C3	0.841	0.797	1.733	1.229	17.498	−6.803	−4.233	28.534	
C3—C4	0.693	0.781	1.473	1.334	23.101	−7.888	−7.104	38.093	
S5—C4	1.000	0.770	1.770	1.044	13.713	−5.216	−3.684	22.613	
C4—H41	0.710	0.443	1.153	1.833	13.452	−19.462	−18.947	51.861	
C4—H42	0.691	0.329	1.019	1.849	16.262	−19.357	−16.635	52.254	
C1—C1′	0.688	0.688	1.375	1.740	27.680	−12.136	−12.009	51.825	
C3—C3′	0.674	0.674	1.348	1.725	33.685	−11.259	−10.632	55.576	
S2—S2′	1.661	1.671	3.330	0.196	1.563	−0.185	−0.073	1.821	
N10—C9	0.582	0.577	1.159	2.404	49.763	−27.129	−20.796	97.688	
C8—C9	0.712	0.706	1.418	1.498	23.766	−10.644	−9.795	44.205	
C7—C8	0.702	0.698	1.400	1.632	25.151	−11.865	−11.017	48.033	
C7—C6	0.720	0.717	1.436	1.382	25.967	−8.758	−8.243	42.968	
C6—H6	0.680	0.397	1.100	1.625	23.063	−19.176	−16.082	58.321	

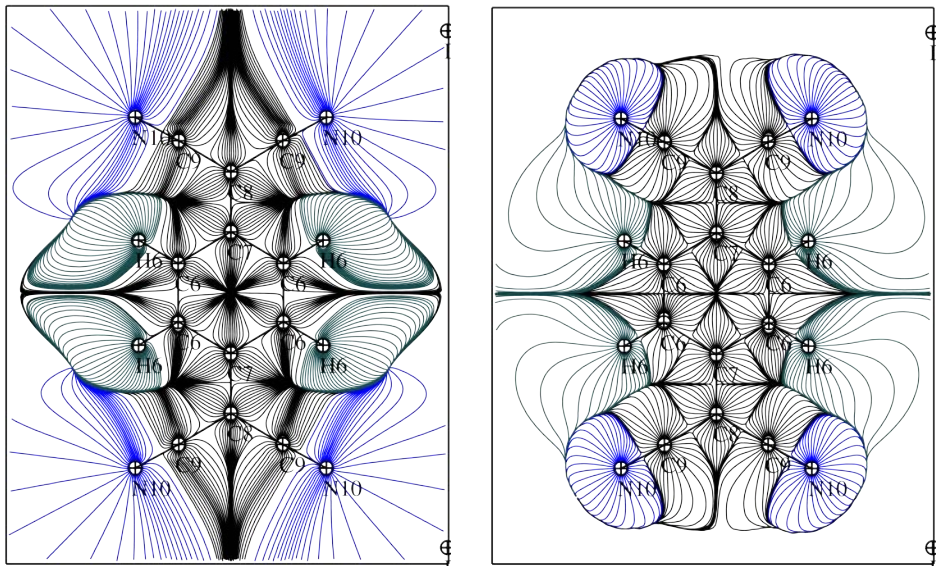


Figure 42: TCNQ: Density (left) and ESP (right) gradient lines. All the atoms are in the plane. Electrostatic Potential field lines are enclosed delimiting a zero charge region.

hole description resembles correctly the sulfur interactions. The statement that sulfurs have sp^3 hybridization is reinforced. If we apply symmetry and rotate the molecule around the Cl-Cl′ bond until we permute internal sulfurs the S2-H interaction points now upside. There is a zigzag interaction between cation and anion stacks connecting S2 lone pairs and hydrogens.

Density values of frontal interactions are greater than those of lateral interactions, table 14. The opposite happens for Laplacian values. Frontal interactions are exceptional, nearly degenerated as its nearly vanishing curvature indicates. The overall of lateral interactions stabilize more the structure.

As well as before, an only electrostatic interaction happens between distant atoms. The ESP is conditionally convergent with the term $\frac{1}{\mathbf{r}}$ (monopoles) converging more smoothly than the ED, $\approx e^{-\mathbf{r}}$.

There is a substantial difference between nitrogen electron density and electrostatic potential basins, figure 47. The integration of charge inside of the ED but outside of the ESP is a partial negative charge. Therefore, most of the anion net charge is concentrated in nitrogens. They are the main charge acceptors. Where does that charge come from?

There is no big difference between sulfur basins. Besides we cannot assume that external sulfurs donate most of the charge. The sulfur ρ-basin total charge is close to zero. Hydrogens and internal sulfurs might play a more important role in charge transfer than expected.

In addition to interactions involving molecules of different stacks there are also π–π interactions along the chain of molecules in the columns, figure 48,47.

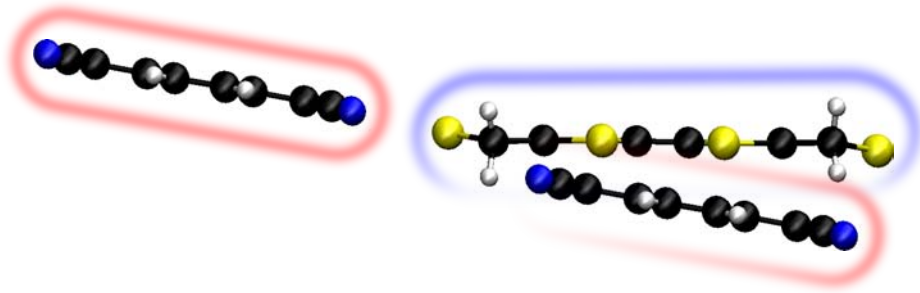


Figure 43: The environment of a BTDMTTF cation: (a) frontal anion and (b) lateral anion constitute the LONGITUDINAL and LATITUDINAL dimers respectively.

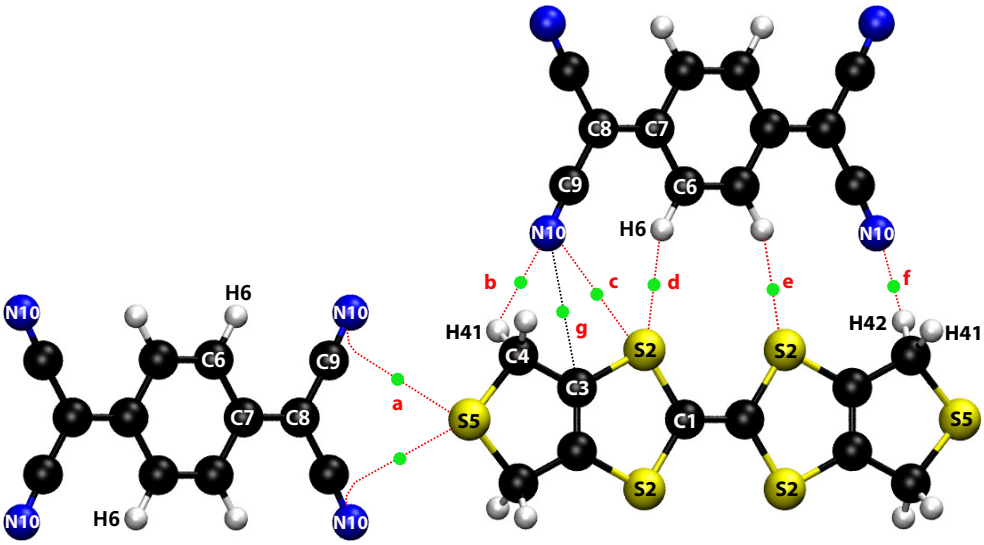


Figure 44: Frontal (a) and lateral (b-g) interactions. There is an only electrostatic interaction (g).

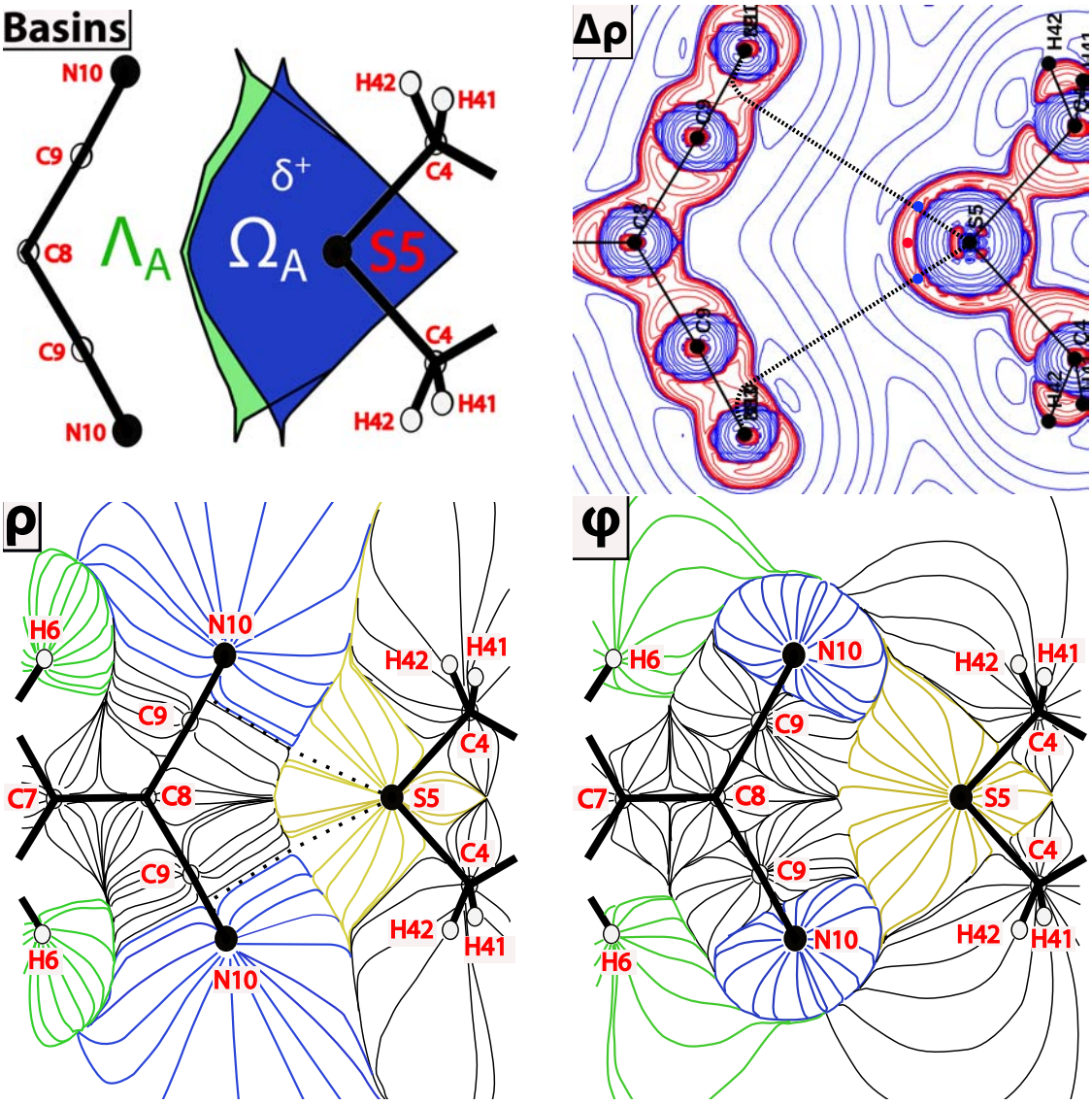


Figure 45: LONGITUDINAL: Density (left) and ESP (middle) gradient lines. The volume of nitrogen density basins is greater than ESP basins. Nitrogens accumulate a negative charge. S5 basin volume is almost not changed. The S5–triple bond path goes between N and C9 basins. Nearly degenerated. Plane of S5 and two N10.

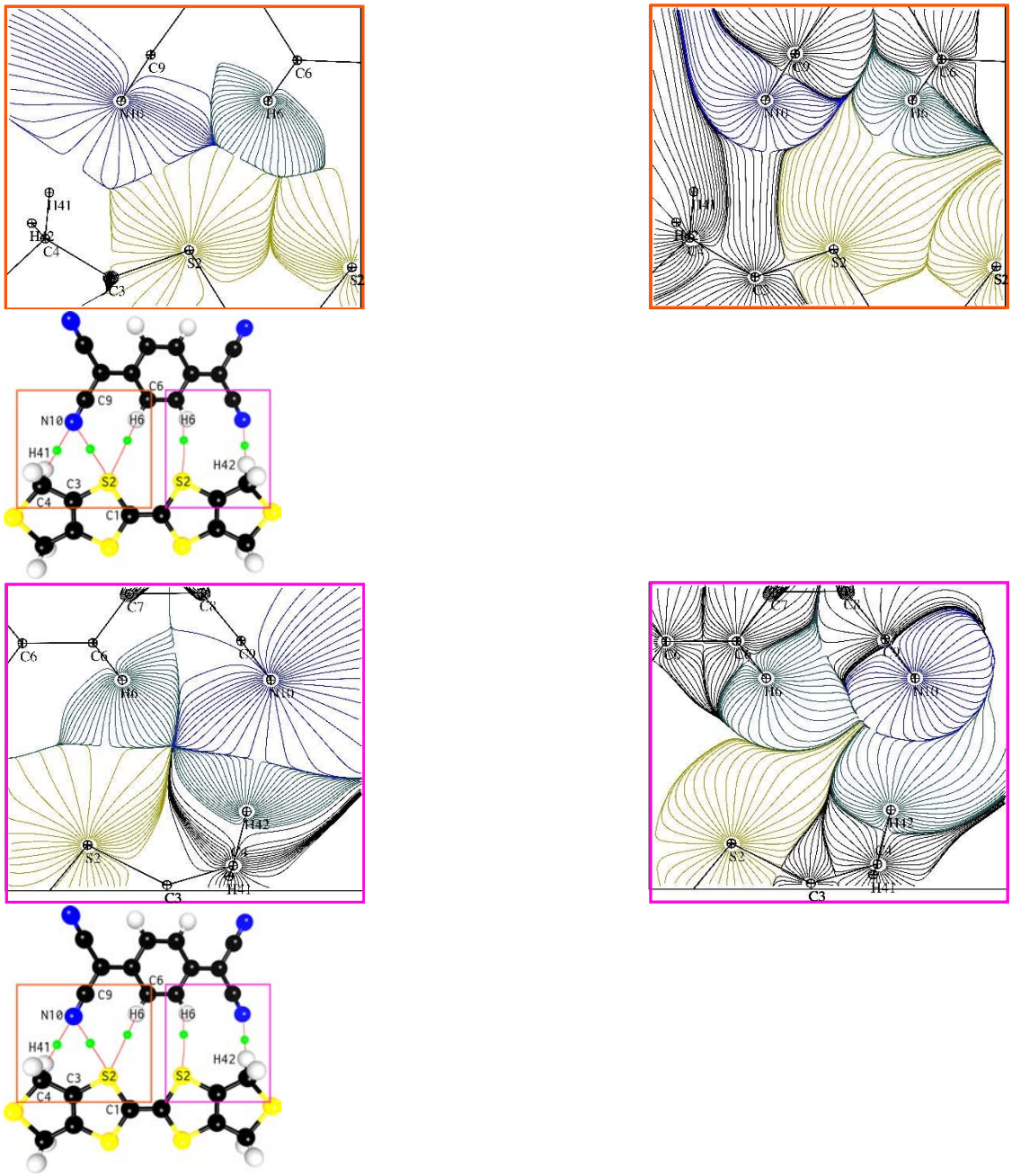


Figure 46: LATITUDINAL: Density and ESP gradient lines. Top images represent the plane N10-H6-S2. Bottom images represent the plane N10-H41-S2.

Table 14: Properties of (3, −1) critical points in the LONGITUDINAL and the LATITUDINAL dimer. The meaning of variables is the same as set for table 13. The value of the Laplacian, $\nabla^2\rho(e/Å^5)$ and $\nabla^2\phi(e/Å^3)$, for intermolecular interactions can be included due to a lower density variability with small position modifications.

ρ: A ··· B	d_A(Å)	d_B(Å)	d_{AB}(Å)	$\rho(e/Å^3)$	$\nabla^2\rho$	λ_1	λ_2	λ_3
C9≡N10 ··· S5	1.739	1.761	3.439	0.402	0.534	−0.090	−0.055	0.679
H41 ··· N10	1.056	1.466	2.522	0.062	0.949	−0.206	−0.163	1.318
S2 ··· N10	1.686	1.528	3.214	0.049	0.645	−0.125	−0.113	0.884
S2 ··· H6	2.057	1.285	3.342	0.013	0.232	−0.030	−0.020	0.283
S2′ ··· H6′	1.965	1.350	3.315	0.020	0.262	−0.055	−0.042	0.359
H42 ··· N10	1.022	1.571	2.593	0.030	0.609	−0.105	−0.098	0.812
ϕ: A ··· B	d_A(Å)	d_B(Å)	d_{AB}(Å)	$\phi(e/Å)$	$\nabla^2\phi$	γ_1	γ_2	γ_3
C9≡N10 ··· S5	1.574	1.860	3.390	0.031	0.460 7	−0.191	−0.107	0.758
C3 ··· N10	2.537	0.988	3.525	1.639	3.625	−0.887	−0.755	5.267
S2 ··· H6	2.330	1.021	3.351	1.248	0.231	−0.491	−0.417	1.139

Sulfurs are given to establish many weak connections with other sulfurs (see table 15). Most of them are between equivalent atoms, therefore they are not relevant to charge transfer. They are not traditional VSCD to VSCC.

7.4.3 Theoretical-experimental correspondence

We start considering (3,-1) intramolecular critical points (see table 16). The theoretical and experimental values are in good correspondence, however there are some nuances that require our attention. The electrostatic potential can be determined by integration of the Poisson equation up to a constant. Therefore, the experimental and theoretical fields should match if we shift one of them taking as a reference one point (e.g. the (3,-1) point between two C1 atoms). Nevertheless we found that the difference is not zero. Surprisingly we obtain a linear dependence between the value of the field (theoretical) at critical points and the difference between both fields (see figure 49).

The linear relationship suggests us that everything is fine but the parameters of the equation need explanation. There is a key difference between both fields: the theoretical potential has been obtained using pseudopotentials. The shape of the theoretical field near nuclei positions would not need to resemble that of the experimental. Instead the values inside the pseudo-core radius should approach the interstitial field near the pseudopotential radius. We could guess that they bring closer following a linear dependence. Thus the parameters that we obtained are those that relate the total and pseudo electrostatic potential. The reason for this explanation is that the intramolecular critical points are located at points close to the pseudopotential radii. The potential is highly influenced by nuclei contributions to the total charge distribution. The value of the field at those points falls in a range of positive values.

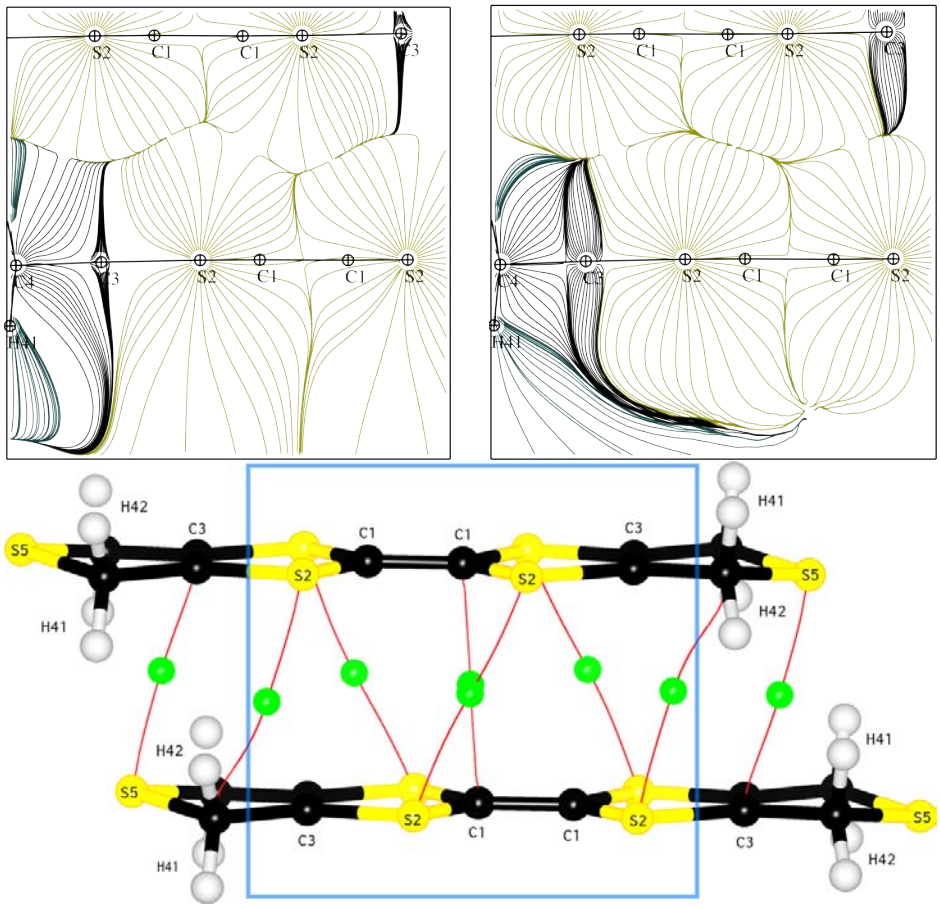


Figure 47: BTDMTTF stack: Density and ESP gradient lines. Planes defined by 4 S2 atoms.

Table 15: BTDMTTF and TCNQ stacks: Density and ESP (3, −1) critical points. The meaning of variables is the same as set for table 13.

ρ: A ⋯ B	d_A(Å)	d_B(Å)	d_{AB}(Å)	ρ(e/Å^3)	$\nabla^2\rho$	λ_1	λ_2	λ_3
S2 ⋯ S2′	1.944	1.944	3.888	0.035	0.297	−0.074	−0.019	0.391
S2 ⋯ S2″	1.948	1.931	3.879	0.035	0.324	−0.056	−0.034	0.414
S2 ⋯ C4	1.925	1.867	3.792	0.037	0.329	−0.081	−0.025	0.435
C1 ⋯ C1	1.757	1.757	3.514	0.036	0.352	−0.033	−0.023	0.407
S5 ⋯ C3-C3	1.803	1.671	3.474	0.045	0.413	−0.091	−0.032	0.536
C7 ⋯ C7	1.649	1.649	3.298	0.047	0.478	−0.062	−0.016	0.555
C8 ⋯ C7	1.640	1.639	3.279	0.048	0.501	−0.071	−0.022	0.595
ϕ: A ⋯ B	d_A(Å)	d_B(Å)	d_{AB}(Å)	ϕ(e/Å)	$\nabla^2\phi$	γ_1	γ_2	γ_3
H41 ⋯ H42	1.190	1.268	2.458	0.424	0.230	−0.345	−0.302	0.877
S2 ⋯ H42	1.800	1.466	3.266	0.342	0.186	−0.219	−0.140	0.545
S2 ⋯ S2′	1.997	1.998	3.995	0.193	0.159	−0.107	−0.046	0.311

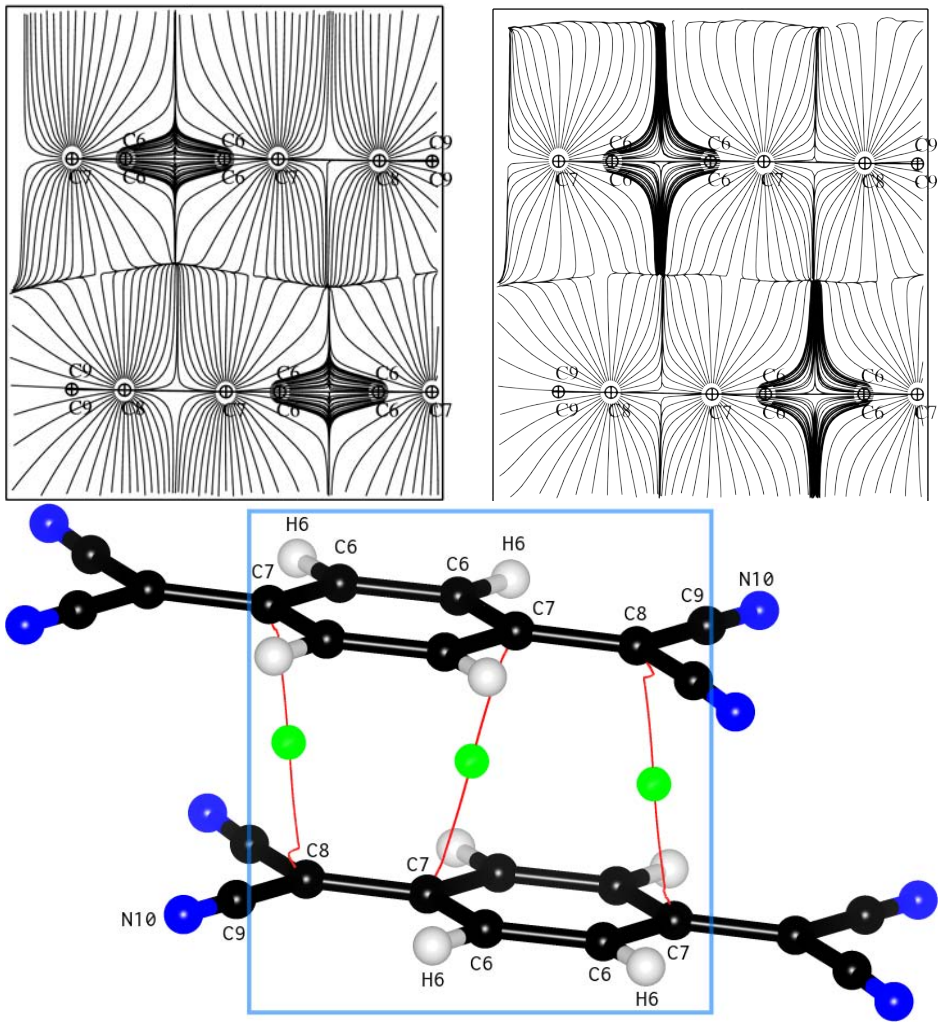


Figure 48: TCNQ stack: Density and ESP gradient lines. Planes defined by C7 and C8 atoms.

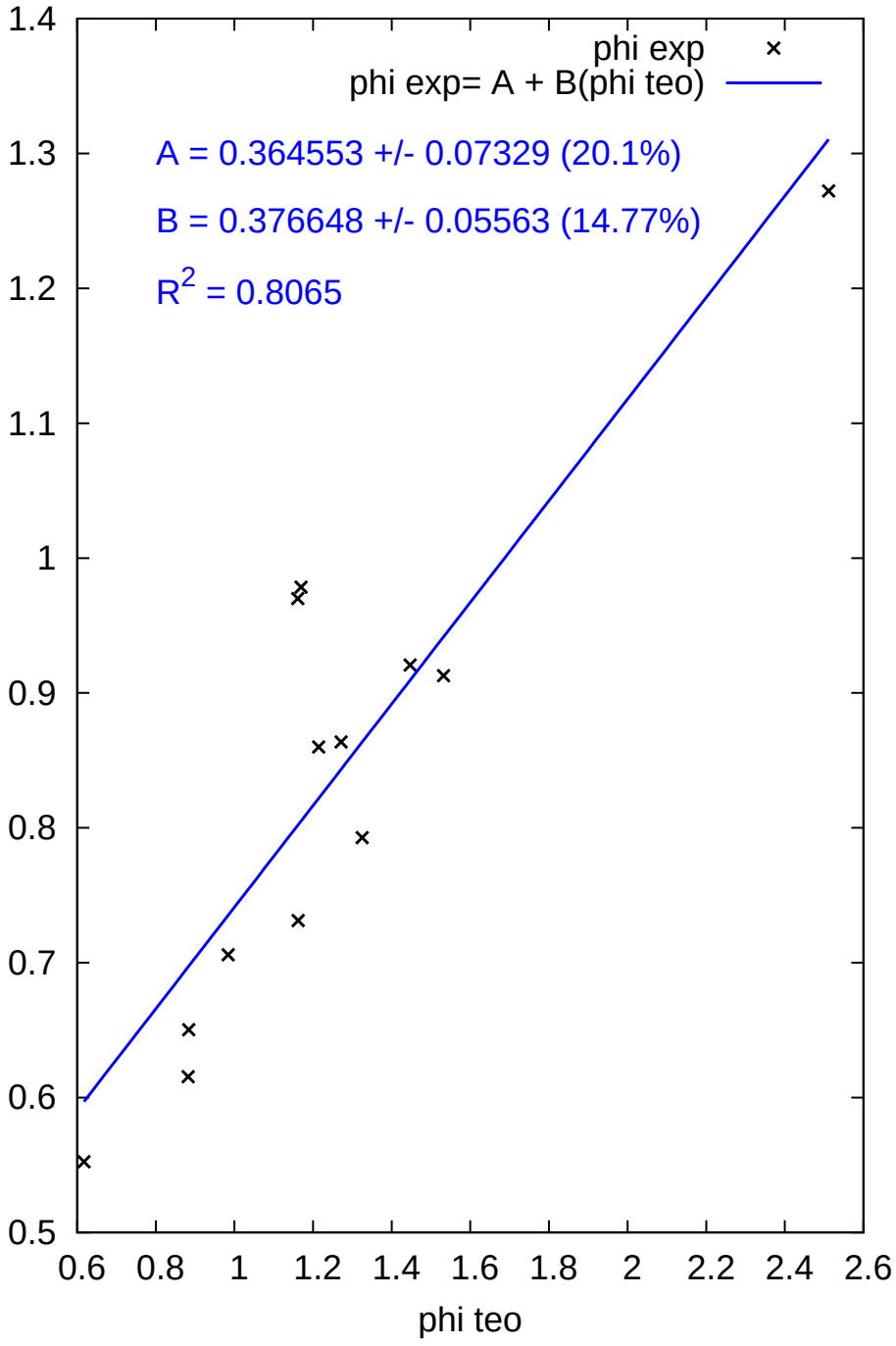


Figure 49: Experimental and theoretical given values exhibit a linear relation at points where the potential is positive.

Table 16: The following is a comprehensive list of ab initio and experimental values of the electrostatic potential ϕ [e/bohr] and laplacian $L(\mathbf{r})$ [e/bohr^3] at the critical points of the potential. Along are attached the their differences. The locations electrostatic potential critical points are analogous to bond critical points for intramolecular interactions.

		ϕ_{opt}	ϕ_{exp}	$\Delta\phi$	$L_{opt}(\mathbf{r})$	$L_{exp}(\mathbf{r})$	$\Delta L(\mathbf{r})$
S5	C4	0.6170	0.1381	-0.4789	1.05	0.5080	0.539
C1	S2	0.8820	0.1538	-0.7282	6.79	0.6544	6.133
C3	S2	0.8839	0.1625	-0.7213	5.17	0.6482	4.522
C3	C4	0.9842	0.1764	-0.8077	5.78	0.8558	4.926
		0.9842			5.79		
C4	H41	1.1608	0.2424	-0.9183	2.70	0.4983	2.196
		1.1608			3.39		
C6	C7	1.1620	0.1828	-0.9791	7.10	0.9619	6.134
		1.1620			7.32		
C4	H42	1.1692	0.2446	-0.9246	2.99	0.6024	2.389
		1.1693			3.80		
C6	H6	1.2144	0.2149	-0.9994	2.28	0.8543	1.427
		1.2144			2.28		
C7	C8	1.2714	0.2159	-1.0555	7.87	0.9317	6.934
C8	C9	1.3246	0.1981	-1.1264	6.83	0.8804	5.950
		1.3247			6.53		
C1	C1	1.4468	0.2301	-1.2166	4.99	1.0254	3.968
C3	C3	1.5320	0.2282	-1.3038	7.90	1.2479	6.652
C9	N10	2.5112	0.3180	-2.1932	1.93	1.8435	0.085
		2.5113			-0.51		
S2	S2		0.0259	0.0259		0.0579	-0.0579
C6	C6						

Another reason for the discrepancies may be rooted in the different structures, pointed out in the methodology, used for experimental and theoretical analyses.

Both theoretical and experimental electrostatic potentials are qualitatively equivalent as shown by figure 50. The critical points that appear for the experimental potential also appear for the theoretical field. The mismatching corresponds to a scaling factor.

7.4.4 Estimation of charge transfered

We want to emphasize that both experimental and ab initio calculations agree on the degree of charge transfer: 0.6 e. The plane wave calculations fall in between the neutral and ionic state for most atoms. All theoretical values lie very close together whereas the values of experimental atomic charges are far

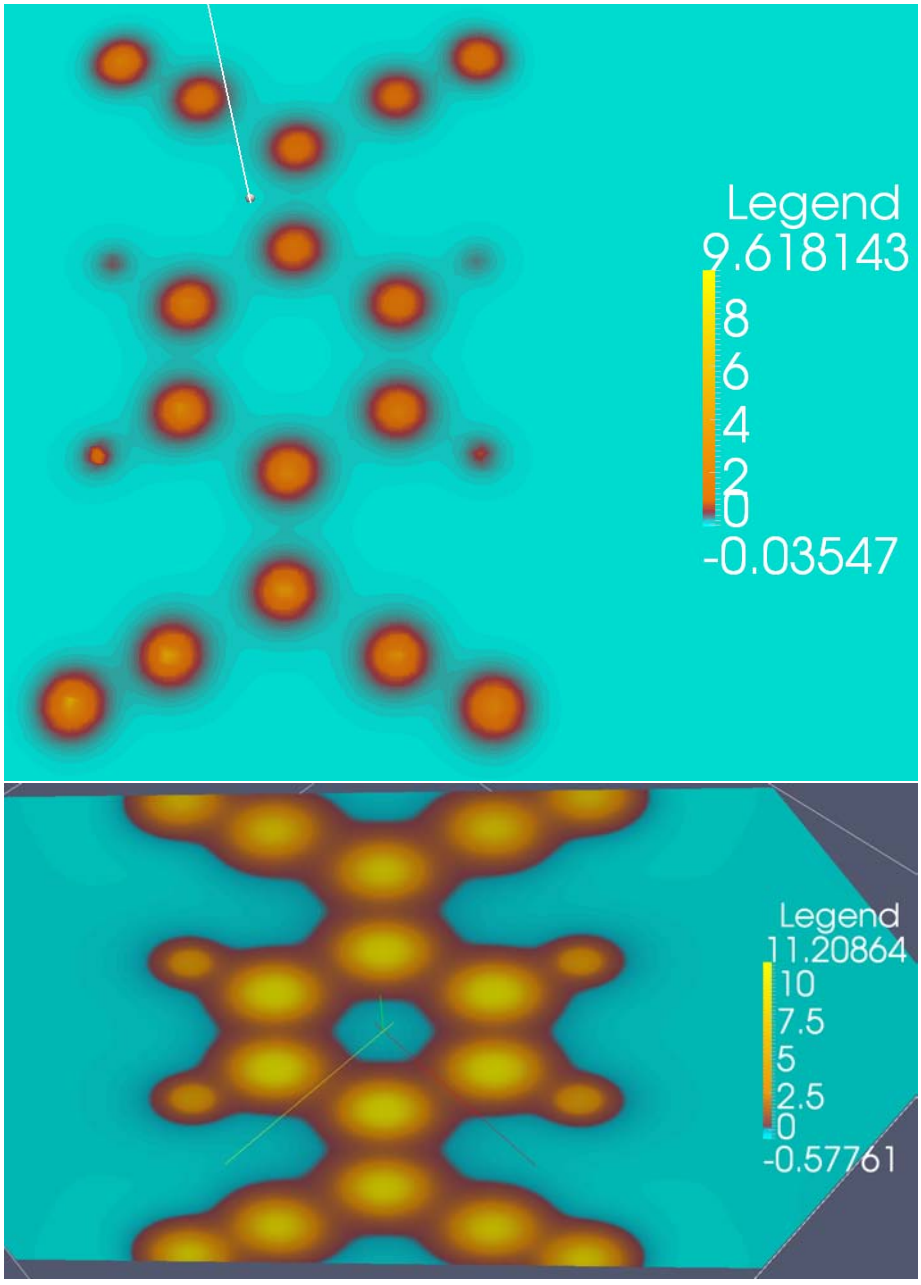


Figure 50: The electrostatic potential map in the plane that contains TCNQ of experimental (top) and theoretical origin (bottom). Even if it is not clearly seen from the figure, there are hollows in front of nitrogen atoms (that correspond to (3,+3) points) in both fields.

apart. That is an error inherent to the multipolar model, that has been designed to fit the overall/whole charge of the system but not to atomic contributions.

A previous work on the system obtained net atomic charges from kappa and multipolar model parameters. The results thereof [47] reported an estimate of the charge transferred upon excitation (q_{CT}) around 0.7 e. Here we report atomic charges integrated for the first time using QTAIM partitioning, a intermolecular charge transfer of 0.62 electron charges per molecule, taking as a reference the neutral state. This value is very similar to bulk TCNQ-TTF, were it is 0.59 e [3]. Ionicity has slightly increased as a result of the inclusion of the additional BTDM cycle.

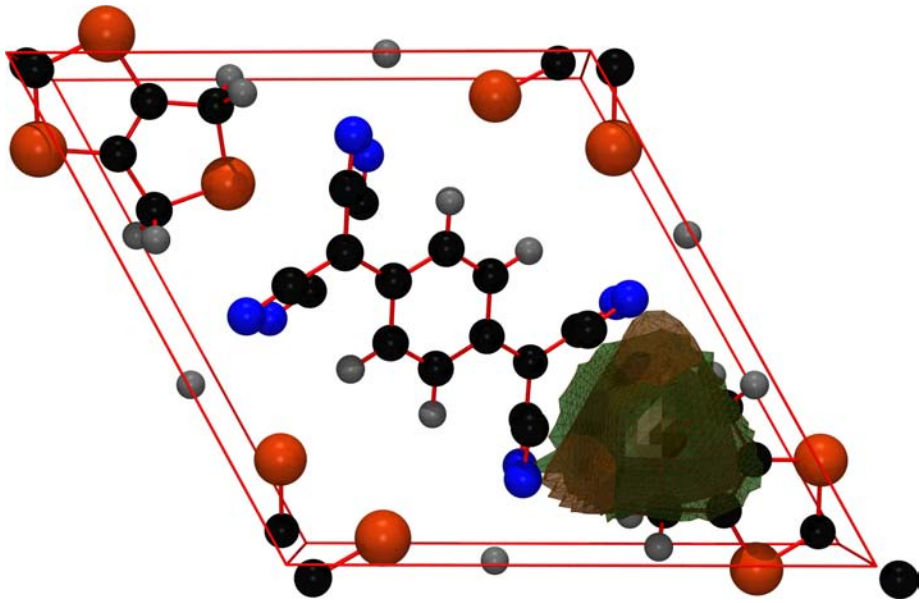


Figure 51: The density (green) and electrostatic potential (red) basins of S5. Hydrogen atoms are grey, carbon atoms are black, nigrogen atoms are blue, and sulphur atoms are orange-red.

All previous pseudo-isolated molecules share a common fact: some basins are open. We have to handle this fact if we are interested in integrating atomic partial and total charges without neglecting the integration far from the nuclei. Taking a cluster, the basins of the inner molecule are finite because it is completely surrounded.

Anions and cations are stacked as columns along the c axis. Each anion or cation is surrounded by 8 counterpart ions in the ab crystallographic planes and two equals along the c direction.

The left side of the table 17 are properties of the cation and the right side are properties of the anion. The integrated charges of molecules are in close agreement with multipolar (Q') and kappa (Q'') populations. The agreement vanishes when we compare atomic charges. For example, C9 and N10 of car-

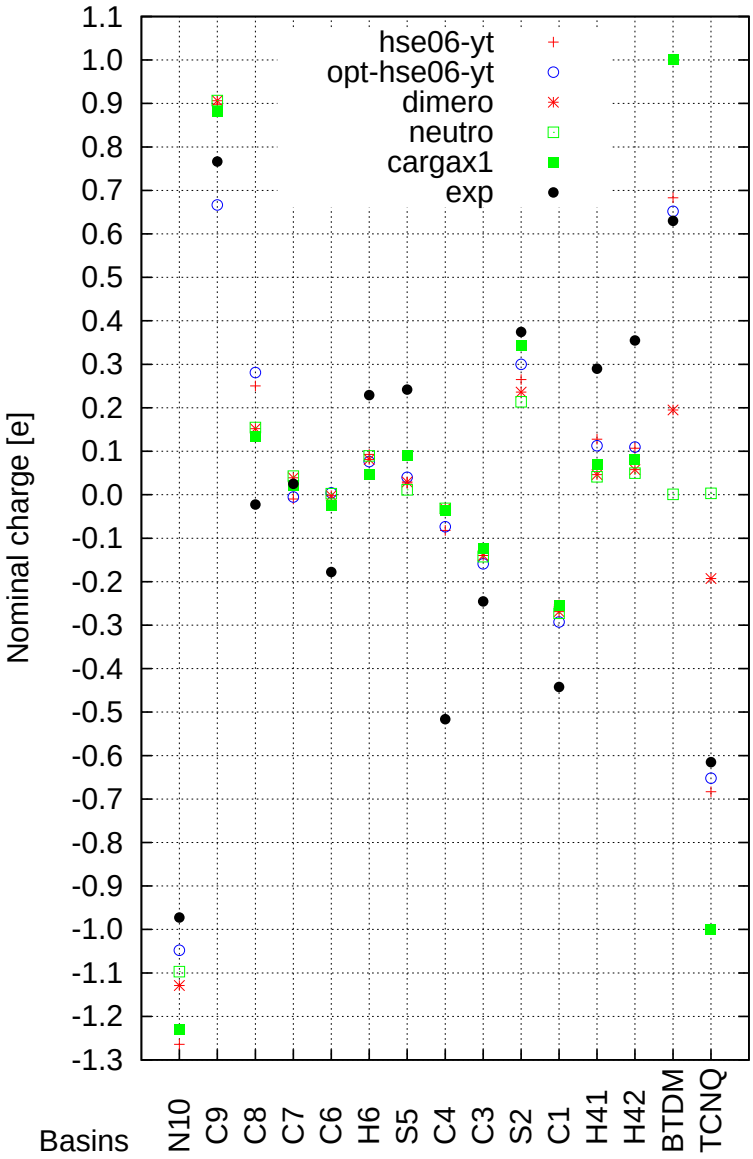


Figure 52: Atomic charges of BTDMTTF and TCNQ integrated in different environments and electronic configurations. `exp` is the experimental density of the crystal, `neutro` are the densities of BTDMTTF and TCNQ separated and in isolation, `dimero` is the density of the dimer BTDMTTF-TCNQ, `hse06-yt` is the bulk density calculated with HSE06 functional without optimization and `opt-hse06-yt` optimized.

bonyl groups. However, refined charges are good for chemical groups (e.g. CN), or molecules.

Nitrogens have a partial charge close to 1 e, but most of the charge is transfered from the contiguous carbon (C9). Sulfurs and hydrogens of BTDMTTF contribute equally to charge donation.

The negative charge of N10 is located inside a large volume whereas the positive charge of internal sulfurs is located in a very small volume. As an average the density is high inside that region of space.

The integration of the density in the ESP basins is good. We loss 2 e in the integration. It is about 0.008 e/basin.

The charge transfer from the integration is 0.63 whereas the values from two multipolar refinements published beforehand are close and suggest us that for chemical groups those parameters estimate right values. If we take a look at both the partial atomic charges and the volume difference between the density and ESP basins we can estimate the localization partial charges. Thus for the S2 atom we can see that the charge is highly concentrated in a small region of space.

S5 basins volume are scarcely different as we expected from plane sections outlined before.

The charge inside ESP basins is 0 (minor numerical errors) as it should be.

Table 17: ED integration inside density basins (Ω) (top block) and ESP basins (Λ) (bottom block) of the inner molecule of the BTDMTTF cluster (left) and TCNQ cluster (right). Q' and Q'' taken from [46]. * Total refers to the whole molecule (BTDMTTF or TCNQ).

A	Ω_A(Å^3)	Q_{Ω_A}(e)	Q'(e)	Q''(e)	A	Ω_A(Å^3)	Q_{Ω_A}(e)	Q'(e)	Q''(e)
S2	23.263	0.375	0.365	0.356	N10	25.285	−0.973	−0.139	−0.135
C3	9.608	−0.245	−0.171	−0.178	C6	11.537	−0.179	−0.045	−0.039
C4	12.841	−0.516	−0.500	−0.512	C7	8.821	0.026	0.088	0.092
H41	5.060	0.288	0.251	0.259	C8	9.598	−0.020	−0.074	−0.066
H42	5.272	0.358	0.300	0.301	C9	4.613	0.767	−0.075	−0.071
S5	21.961	0.241	0.328	0.336	H6	6.105	0.230	0.045	−0.039
C1	10.616	−0.442	−0.444	−0.454					
Total*	289.33	0.63	0.75	0.67	Total*	227.00	−0.62	−0.75	−0.67
A	Λ_A(Å^3)	Q_{Λ_A}(e)			A	Λ_A(Å^3)	Q_{Λ_A}(e)		
S2	23.691	0.002			N10	8.749	0.006		
C3	7.803	0.003			C6	10.551	−0.006		
C4	8.750	−0.004			C7	7.926	0.010		
H41	10.888	0.009			C8	8.412	−0.009		
H42	12.890	−0.005			C9	13.416	−0.005		
S5	25.073	0.005			H6	11.074	0.014		
C1	6.803	0.006							
Total*	319.84	0.042			Total*	207.84	0.038		

There is a huge polarization of the C9–N10 bond. Curiously, there is no Laplacian zero isosurface between them. The charge depletion along the bond path at the BCP is very high but the charge concentration along the other two directions is also big. The charge concentration between the two atoms is hight but inside the N basin.

The charge is polarized but the IAS is very close to the C9. The value of the curvatures is very high.

Integration of the density — I used the valence — inside the electrostatic potential basins is 0 within a maximum error of 0.07 e.

7.5 CONCLUSIONS

We have unveiled all the interactions that stabilize the BTDMTTF-TCNQ crystal using the topology of the density and electrostatic potential fields. Also we have tried to infer what is the role of those interactions in the charge transfer process. We have found that the shape of the basins of the electrostatic potential and the density are very similar, that allows us to locate where are the partial charges of the atom. Therefore we attempted to unravel how charge transference occurs between the two molecules building upon the intersection of both basins so we concluded that the atoms labeled as S5 are not the only ones of

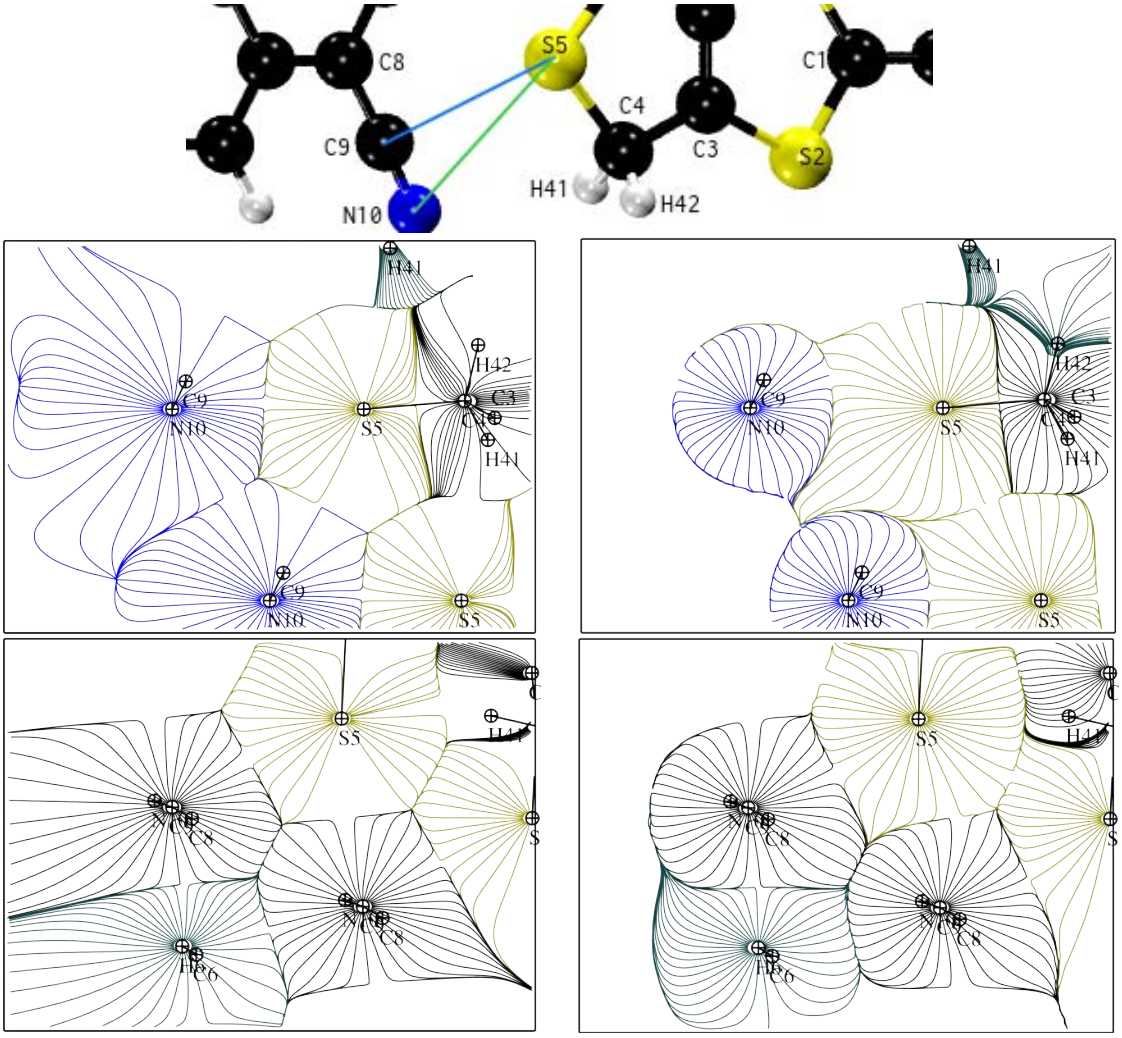


Figure 53: Interactions of S5 with anion atoms. Top planes defined by S5 and two N10 atoms. Bottom planes defined by S5 and two C9 atoms.

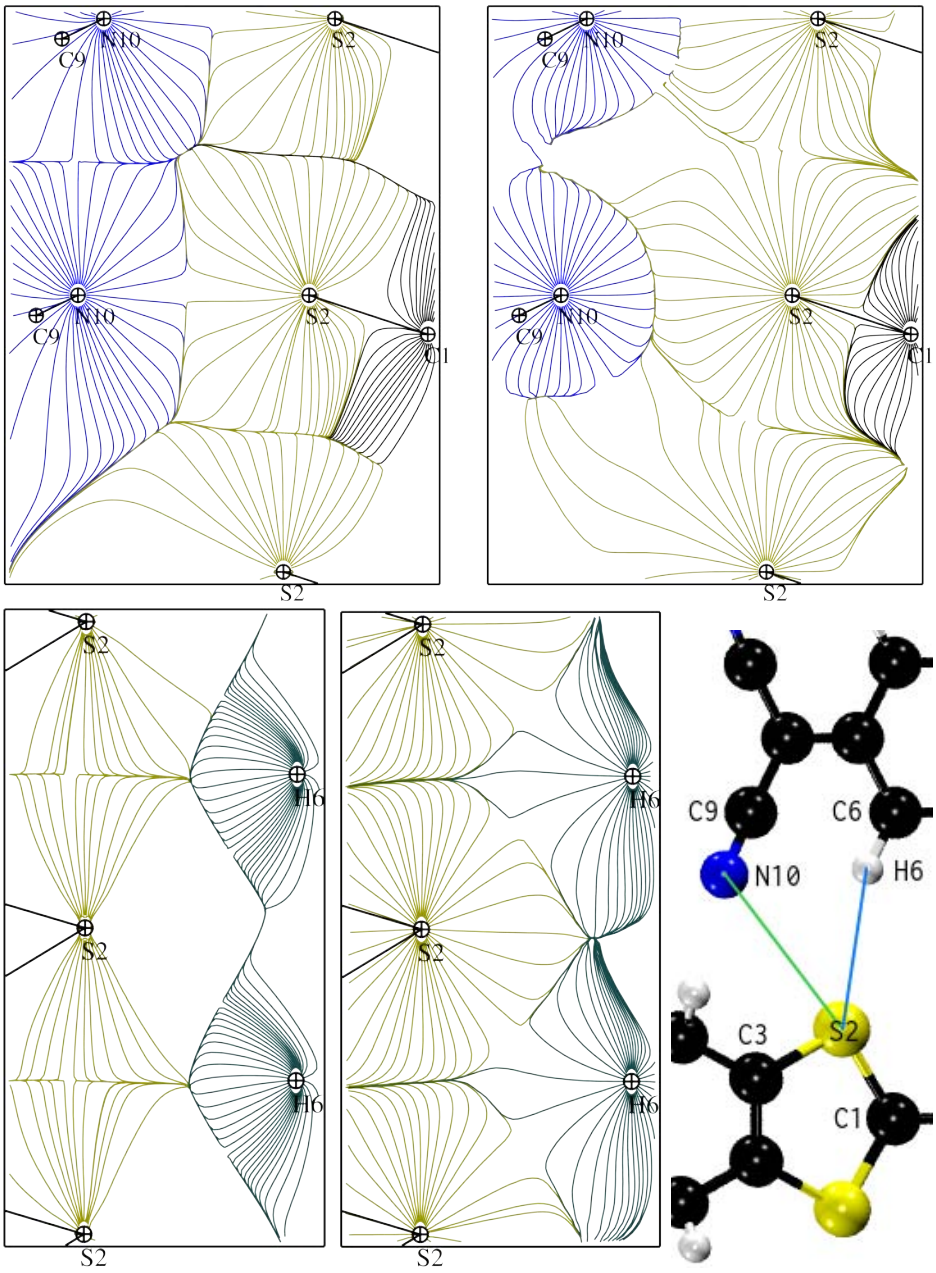


Figure 54: Interactions of S2 with anion atoms. Planes coplanar to S2 and N10 atoms; planes coplanar to S2 and H6 atoms.

the BTDMTTF molecule participating, but that also the hydrogens (with labels H41 and H42) have a non negligible contribution. However, the total charge transferred that is obtained from integration in QTAIM basins is not distant from previous estimations. That says us that even if less than one electron is transfered per molecule the system is a good conductor. On the other hand, our attempts to bring together the topology of the experimental and theoretical electrostatic potential have not been as productive as we expected in part due to the sensibility of the electrostatic potential to changes in the density. Given that the type of functions used to approximate the density in both models are different we have not been capable of going further.

Part IV

Conclusions

8 CONCLUSIONES

A continuación se procede a presentar un compendio de las conclusiones extraídas en esta tesis.

En el capítulo 4 hemos visto que las energías de enlace de moléculas poliatómicas no se pueden obtener a menos que se realicen aproximaciones y que incluso en el caso de moléculas diatómicas necesitamos fijar un estado de referencia. Fijando esa referencia se puede obtener una medida de la fuerza del enlace, pero de nuevo no tenemos a nuestro alcance un método para llevar a cabo dicha medida. Es aquí donde la partición energética IQA juega un papel clave para sentar sobre una base firme el concepto de energía intrínseca de enlace. Junto con el resto de herramientas de QCT podemos caracterizar la referencia, i.e. identificar con que estado electrónico se corresponde más fidedignamente el estado de valencia. Se deduce de los resultados que las configuraciones electrónicas de alto espín de los fragmentos estudiados tienen un peso menor del que comúnmente se les atribuye. Esto nos indica que deberíamos ser especialmente cautos a la hora de asignar estados de valencia sin una justificación previa razonable. Una vez que esto se ha tenido en cuenta no hay grandes cambios. Asimismo, las energías de relajación de los hidrógenos en el metano son importantes. También se ha observado claramente como los componentes sigma y pi de del enlace triple $N \equiv N$ se forman en dos pasos.

En MULTIPOLAR EXPANSION OF THE EXCHANGE–CORRELATION INTERACTION ENERGY hemos puesto de relieve que las energías de intercambio y correlación interatómicas que se utilizan en las teorías de enlace químico en el epacio real, las cuales contabilizan la contribución covalente a una interacción interatómica dada, pueden ser aproximadas por una expansión multipolar convencional. Rigurosamente, la serie diverge cuando los átomos están directamente enlazados, aunque puede entenderse como convergente asintóticamente. El truncado de las series hasta $l_1 + l_2 = 2$ (incluyendo hasta interacciones carga-cuadrupolo) tiende a dar resultados que son precisos hasta un pequeño porcentaje en interacciones 1-n, con $n > 2$, e incluso hasta un 10% en muchos casos directamente enlazados o 1-2. En el caso de $n > 2$ la serie converge en muchos casos, e incluyendo términos extra provee de precisión extra. Al contrario, la consideración de contribuciones l mayores en interacciones 1-2 tiende a deteriorar seriamente los resultados. Dado que la carga necesaria para calcular la serie multipolar es considerablemente menor que la de la expansión bipolar exacta, nuestros resultados podrían ser importantes para estimar interacciones covalentes en esos casos en los que la integración exacta no es viable. También podría ser usado para reducir el coste computacional en descomposiciones energéticas con el método IQA de sistemas grandes, donde muchos términos con

valores pequeños, asociados a interacciones de largo alcance y tremendamente costosos computacionalmente, pueden ser aproximados ahora sin el riesgo de tener una disminución de precisión.

Resumiendo el capítulo 6, hemos demostrado que existe una conexión notable entre la teoría de Kohn-Resta para estados aislantes a través del tensor de localización (o en el caso de sistemas moleculares su tensor de extensión total) y los bien conocidos indicadores que se utilizan en la teoría de enlace químico como ordenes de enlace. Esto se ha alcanzado con un particionado del tensor de localización en componentes intra- e interatómicos. Una forma de hacerlo respetando la invarianza frente al cambio de orbitales se empieza escribiendo el tensor de localización (TL) en términos de las matrices de densidad reducidas y luego particionando el espacio en regiones atómicas de acuerdo con la topología química mecanocuántica. La convergencia o divergencia del TL en el límite termodinámico, asociado a las propiedades conductoras o aislantes, depende exclusivamente en el ritmo de decaimiento de sus componentes interatómicos. Este último es denominado el indice de delocalización químico, una forma moderna de entender el orden de enlace. De esta forma la química de los estados fundamentales y la física de la conductividad se encuentran entrelazados. Esperamos que esta conexión sea útil en la búsqueda y diseño de conductores, o aislantes, de dimensión reducida ya que el tensor total TL puede ser escrito como una suma de componentes atómicos (o grupos funcionales). Cada átomo o grupo funcional en el sistema podría ser clasificado como esencial si su contribución al tensor diverge, o secundario si converge, si consideramos la conductividad eléctrica. Finalmente, esperamos que esta categorización pueda ser usada ventajosamente en el diseño racional de nuevos materiales.

En el capítulo TOPOLOGY OF THE ELECTROSTATIC POTENTIAL IN SOLIDS se han desvelado todas las interacciones que estabilizan el crystal BTDMTTF-TCNQ usando la topología de la densidad y el potencial electrostático. También se ha intentado inferir que papel juegan en la transferencia de carga. Nos hemos fijado en que no existe mucha diferencia entre las cuencas del potencial y la densidad, lo que nos permite asignar a su intersección el significado de carga parcial. Intentamos desentrañar como es el proceso de reorganización de carga entre ambas moléculas usando esas intersecciones y llegamos a la conclusión de que los átomos etiquetados como S5 no son los únicos que contribuyen a la transferencia de energía, sino que los hidrógenos (con etiquetas H41 y H42) también forman parte del proceso de transferencia de carga. Contabilizando la transferencia total entre una molécula y otra a partir de la partición QTAIM hemos obtenido un valor muy similar a las aproximaciones anteriores. Lo cual nos dice que a pesar de ser una tasa menor que un electrón por molécula el sistema es un buen conductor. Los intentos de acercar la descripción teórica y experimental del potencial no ha sido tan fructífera como esperábamos en parte por la sensibilidad del potencial ante cambios de la densidad. Como el tipo de funciones con las que se aproxima la densidad en ambos modelos son diferentes no hemos podido o sabido ir más allá en nuestro estudio.

Part V

Appendix

A ADDITIONAL MATHEMATICAL RELATIONS

CONTENTS

A.1 THE POSITION SPREAD TENSOR (PST) FROM REDUCED DENSITY MATRICES (RDM)

We will examine the mathematical derivations leading to the TPS in real space and details on its implementation in the PROMOLDEN code. Let $\mathbf{r}_i$ the vector position of the i^{th} electron of a N-electron molecule, Ψ the multielectron wavefunction, and

$$\mathbf{R} = \sum_i^N \mathbf{r}_i. \tag{A–1}$$

The total position spread (TPS) tensor is defined as [88, 163]

$$\boldsymbol{\Lambda} = \langle\Psi|\hat{\mathbf{R}}\otimes\hat{\mathbf{R}}|\Psi\rangle - \langle\Psi|\hat{\mathbf{R}}|\Psi\rangle\otimes\langle\Psi|\hat{\mathbf{R}}|\Psi\rangle \equiv \langle\hat{\mathbf{R}}\otimes\hat{\mathbf{R}}\rangle - \langle\hat{\mathbf{R}}\rangle\otimes\langle\hat{\mathbf{R}}\rangle. \tag{A–2}$$

Here, bold fonts are used to indicate vectors or tensors, depending on the context, and the $\otimes$ symbol stands for a tensor or cartesian product. For example, the cartesian components of the $\mathbf{r}\otimes\mathbf{r}$ tensor are $(\mathbf{r}\otimes\mathbf{r})_{ab} = x_a x_b$, and the six independent components of the symmetric $\boldsymbol{\Lambda}$ are

$$\Lambda_{ab} = \Lambda_{ba} = \langle X_a X_b\rangle - \langle X_a\rangle\langle X_b\rangle, \qquad a,b = x,y,z. \tag{A–3}$$

$\boldsymbol{\Lambda}$ is a cumulant and size extensive.

Using equation (A–1) and $\hat{\mathbf{R}}\otimes\hat{\mathbf{R}} = \sum_{i\neq j}^N \hat{\mathbf{r}}_i\otimes\hat{\mathbf{r}}_j + \sum_i^N \hat{\mathbf{r}}_i\otimes\hat{\mathbf{r}}_i$ in (A–2), and taking into account electron indistinguishability, $\boldsymbol{\Lambda}$ becomes

$$\boldsymbol{\Lambda} = N(N-1)\langle\Psi|\hat{\mathbf{r}}_1\otimes\hat{\mathbf{r}}_2|\Psi\rangle + N\langle\Psi|\hat{\mathbf{r}}_1\otimes\hat{\mathbf{r}}_1|\Psi\rangle - N\langle\Psi|\hat{\mathbf{r}}_1|\Psi\rangle\otimes N\langle\Psi|\hat{\mathbf{r}}_2|\Psi\rangle. \tag{A–4}$$

Now, we use the definition of the first order, $\rho(\mathbf{r}_1)$, and the second order, $\rho_2(\mathbf{r}_1,\mathbf{r}_2)$, spinless densities, which are nothing but the electron density and the pair density, respectively:

$$\begin{aligned}\rho(\mathbf{r}_1) &= N\int d\sigma_1\, dx_2\cdots dx_N \Psi^*\Psi,\\ \rho_2(\mathbf{r}_1,\mathbf{r}_2) &= N(N-1)\int d\sigma_1\, d\sigma_2\, dx_3\cdots dx_N \Psi^*\Psi,\end{aligned} \tag{A–5}$$

where $\mathbf{x}_i = \mathbf{r}_i\sigma_i$ is the space-spin coordinate of electron i. Then,

$$\begin{aligned}\boldsymbol{\Lambda} &= \int d\mathbf{r}_1 d\mathbf{r}_2 (\hat{\mathbf{r}}_1\otimes\hat{\mathbf{r}}_2)\,\rho_2(\mathbf{r}_1,\mathbf{r}_2) + \int d\mathbf{r}_1 (\hat{\mathbf{r}}_1\otimes\hat{\mathbf{r}}_1)\rho(\mathbf{r}_1)\\ &- \int d\mathbf{r}_1 \hat{\mathbf{r}}_1\,\rho(\mathbf{r}_1)\otimes\int d\mathbf{r}_2\,\hat{\mathbf{r}}_2\rho(\mathbf{r}_2).\end{aligned} \tag{A–6}$$

The above equation may be notably simplified by using $\rho_2(\mathbf{r}_1,\mathbf{r}_2) = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) - \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)$, where ρ_{xc} is the exchange-correlation density, the part of the pair density containing all pure quantum mechanical effects. Using this expression

for ρ_2 in (A–6), the part of $\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)$ compensates the third term, and $\mathbf{\Lambda}$ results

$$\mathbf{\Lambda} = -\int d\mathbf{r}_1 d\mathbf{r}_2 (\hat{\mathbf{r}}_1 \otimes \hat{\mathbf{r}}_2)\, \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) + \int d\mathbf{r}_1 (\hat{\mathbf{r}}_1 \otimes \hat{\mathbf{r}}_1)\, \rho(\mathbf{r}_1). \tag{A–7}$$

If we now use the key identity $\rho(\mathbf{r}_1) = \int d\mathbf{r}_2 \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$ in the second integral of (A–7), we obtain

$$\mathbf{\Lambda} = \int d\mathbf{r}_1 d\mathbf{r}_2 \,[\mathbf{r}_1 \otimes (\mathbf{r}_1 - \mathbf{r}_2)]\, \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2), \tag{A–8}$$

$$= \int d\mathbf{r}_1 d\mathbf{r}_2 \,[\mathbf{r}_2 \otimes (\mathbf{r}_2 - \mathbf{r}_1)]\, \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2), \tag{A–9}$$

where the last equation arises as a consequence of the invariance of $\mathbf{\Lambda}$ with respect to the $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ exchange and the equality $\rho_{xc}(\mathbf{r}_2, \mathbf{r}_1) = \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$. Taking the average of (A–8) and (A–9), and defining the interparticle position vector $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$, $\mathbf{\Lambda}$ may be finally written in the following explicitly origin independent form

$$\mathbf{\Lambda} = \frac{1}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \; \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) \;\; (\mathbf{r}_{12} \otimes \mathbf{r}_{12})\,. \tag{A–10}$$

The origin independence of $\mathbf{\Lambda}$ may also be explicitly proven by writing any of its six independent components, $\mathbf{\Lambda}_{ab}$, in a shifted frame, $\mathbf{r}'_i = \mathbf{r}_i + \mathbf{u}$. Using (A–8) one has

$$\Lambda'_{ab} = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \, \rho_{xc} \, (x_{1a} + u_a) \, [x_{1b} - x_{2b}] \tag{A–11}$$

$$= \int d\mathbf{r}_1 \int d\mathbf{r}_2 \, \rho_{xc} \, x_{1a} \, [x_{1b} - x_{2b}] + u_a \int d\mathbf{r}_1 \int d\mathbf{r}_2 \, \rho_{xc} \; [x_{1b} - x_{2b}]\,, \tag{A–12}$$

where we have denoted with a prime ($'$) the position spread tensor in the displaced reference system and obviated the dependence of ρ_{xc} on $\mathbf{r}_1$ and $\mathbf{r}_2$. If the double integrals are written as $\int\int \cdots = \sum_A \sum_B \int_A \int_B \ldots$, the AA term of the second contribution of (A–12) becomes

$$\int_A \int_A \rho_{xc} \, [x_{1b} - x_{2b}] = \int_A d\mathbf{r}_1 x_{1b} \int d\mathbf{r}_2 \rho_{xc} - \int_A d\mathbf{r}_2 x_{2b} \int d\mathbf{r}_1 \rho_{xc} \tag{A–13}$$

$$= \int_A d\mathbf{r}_1 \; x_{1b} \; G^A(\mathbf{r}_1) - \int_A d\mathbf{r}_2 \; x_{2b} \; G^A(\mathbf{r}_2) = 0, \tag{A–14}$$

where

$$G^\Omega(\mathbf{r}_1) = \int_\Omega d\mathbf{r}_2 \; \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) \tag{A–15}$$

is the domain averaged Fermi Hole (DAFH) of the Ω domain [152, 153]. On the other hand the AB and BA terms of the second contribution of (A–12) become

$$\int_A \int_B \rho_{xc} \, [x_{1b} - x_{2b}] = \int_A G^B x_b - \int_B G^A x_{b,}, \tag{A–16}$$

$$\int_B \int_A \rho_{xc} \, [x_{1b} - x_{2b}] = \int_B G^A x_b - \int_A G^B x_b, \tag{A–17}$$

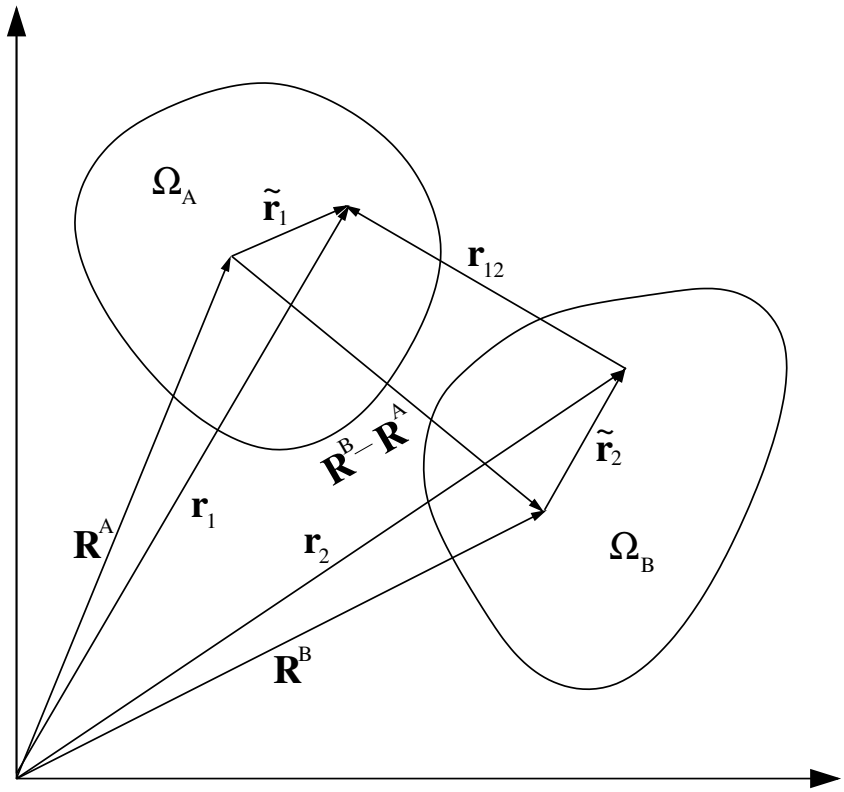


Figure 55: Coordinate System

i.e the AA contribution is zero, and the AB and BA contributions are equal and of opposite sign. Consequently $\int\int \rho_{xc}\,[x_{1b} - x_{2b}] = 0$. This proves that $\Lambda_{ab} = \Lambda'_{ab}$, i.e. that Λ_{ab} is invariant with respect of a translation of the reference system.

A.2 PARTITIONING THE TOTAL POSITION SPREAD TENSOR INTO LOCAL CONTRIBUTIONS

We partition each component Λ_{ab} into local contributions using

$$\Lambda_{ab} = \sum_{A \geqslant B} \Lambda_{ab}^{AB} \tag{A–18}$$

$$\Lambda_{ab}^{AA} = \lambda_{ab}^{AA} \tag{A–19}$$

$$\Lambda_{ab}^{AB} = \lambda_{ab}^{AB} + \lambda_{ab}^{BA} \tag{A–20}$$

$$\lambda_{ab}^{AB} = \int_A d\mathbf{r}_1 \int_B d\mathbf{r}_2\, \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)\, x_{1a}\, [x_{1b} - x_{2b}]\,. \tag{A–21}$$

To compute Λ_{ab}^{AB} it is convenient to transform the coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$, which are referred to a common reference system, to their own local reference frames. Using $\mathbf{r}_1 = \tilde{\mathbf{r}}_1 + \mathbf{R}^A$ and $\mathbf{r}_2 = \tilde{\mathbf{r}}_2 + \mathbf{R}^B$ (See figure 56) or $x_{1a} = \tilde{x}_{1a} + R_a^A$ and $x_{2b} = \tilde{x}_{2b} + R_b^B$, we have

$$\begin{aligned}
\Lambda_{ab}^{AB} &= \int_A d\mathbf{r}_1 \int_B d\mathbf{r}_2 \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)(\tilde{x}_{1a}+R_a^A)(\tilde{x}_{1b}+R_b^A) - (\tilde{x}_{1a}+R_a^A)(\tilde{x}_{2b}+R_b^B) \\
&+ \int_B d\mathbf{r}_1 \int_A d\mathbf{r}_2 \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)(\tilde{x}_{1a}+R_a^B)(\tilde{x}_{1b}+R_b^B) - (\tilde{x}_{1a}+R_a^B)(\tilde{x}_{2b}+R_b^A) \\
&= \langle \tilde{x}_{1a}\tilde{x}_{1b}\rangle_{AB} + \langle \tilde{x}_{1a}R_b^A\rangle_{AB} + \langle R_a^A\tilde{x}_{1b}\rangle_{AB} + \langle R_a^A R_b^A\rangle_{AB} \\
&- \langle \tilde{x}_{1a}\tilde{x}_{2b}\rangle_{AB} - \langle \tilde{x}_{1a}R_b^B\rangle_{AB} - \langle R_a^A\tilde{x}_{2b}\rangle_{AB} - \langle R_a^A R_b^B\rangle_{AB} \\
&+ \langle \tilde{x}_{1a}\tilde{x}_{1b}\rangle_{BA} + \langle \tilde{x}_{1a}R_b^B\rangle_{BA} + \langle R_a^B\tilde{x}_{1b}\rangle_{BA} + \langle R_a^B R_b^B\rangle_{BA} \\
&- \langle \tilde{x}_{1a}\tilde{x}_{2b}\rangle_{BA} - \langle \tilde{x}_{1a}R_b^A\rangle_{BA} - \langle R_a^B\tilde{x}_{2b}\rangle_{BA} - \langle R_a^B R_b^A\rangle_{BA}, \qquad \text{(A–22)}
\end{aligned}$$

where the notation $\langle \hat{o}\rangle_{\Omega\Omega'} \equiv \int_A d\mathbf{r}_1 \int_B d\mathbf{r}_2\ \hat{o}\ \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)$ has been used. Simplifying (A–22)

$$\begin{aligned}
\Lambda_{ab}^{AB} &= \langle \tilde{x}_{1a}\tilde{x}_{1b}\rangle_{AB} - \langle \tilde{x}_{1a}\tilde{x}_{2b}\rangle_{AB} + \langle \tilde{x}_{1a}\tilde{x}_{1b}\rangle_{BA} - \langle \tilde{x}_{1a}\tilde{x}_{2b}\rangle_{BA} \\
&+ \left[R_b^A - R_b^B\right]\left[\langle \tilde{x}_{1a}\rangle_{AB} - \langle \tilde{x}_{1a}\rangle_{BA}\right] + R_a^A\left[\langle \tilde{x}_{1b}\rangle_{AB} - \langle \tilde{x}_{2b}\rangle_{AB}\right] \\
&- R_a^B\left[\langle \tilde{x}_{2b}\rangle_{BA} - \langle \tilde{x}_{1b}\rangle_{BA}\right] + \left[R_a^A - R_a^B\right]\left[R_b^A - R_b^B\right] N_{AB}, \qquad \text{(A–23)}
\end{aligned}$$

where $N_{AB} \equiv \langle\rangle_{AB} \equiv \langle\rangle_{BA} \equiv N_{BA} = \int_A d\mathbf{r}_1 \int_B d\mathbf{r}_2 \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)$ and $N_{AB} + N_{BA} = \delta_{AB}$ is the delocalization index. Taking into account that $\langle \tilde{x}_{2b}\rangle_{BA} = \langle \tilde{x}_{1b}\rangle_{AB}$, $\langle \tilde{x}_{1b}\rangle_{BA} = \langle \tilde{x}_{2b}\rangle_{AB}$, and $\langle \tilde{x}_{1a}\rangle_{BA} = \langle \tilde{x}_{2a}\rangle_{AB}$ (A–23) may also be written in the more symmetrical form [1]

$$\begin{aligned}
\Lambda_{ab}^{AB} &= \langle x_{1a}x_{1b}\rangle_{AB} - \langle x_{1a}x_{2b}\rangle_{AB} + \langle x_{1a}x_{1b}\rangle_{BA} - \langle x_{1a}x_{2b}\rangle_{BA} \\
&- R_b^{AB}\left[\langle x_{1a}\rangle_{AB} - \langle x_{2a}\rangle_{AB}\right] - R_a^{AB}\left[\langle x_{1b}\rangle_{AB} - \langle x_{2b}\rangle_{AB}\right] \\
&+ R_a^{AB}R_b^{AB}N_{AB} \qquad \text{(A–24)}
\end{aligned}$$

where we have defined $\mathbf{R}^{AB} = \mathbf{R}^B - \mathbf{R}^A$.

Performing the same steps with the diagonal element, Λ_{ab}^{AA}, we have

$$\begin{aligned}
\Lambda_{ab}^{AA} &= \int_A d\mathbf{r}_1 \int_A d\mathbf{r}_2 \rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)(x_{1a}+R_a^A)(x_{1b}+R_b^A) - (x_{1a}+R_a^A)(x_{2b}+R_b^A) \\
&= \langle x_{1a}x_{1b}\rangle_{AA} + \langle x_{1a}R_b^A\rangle_{AA} + \langle R_a^A x_{1b}\rangle_{AA} + \langle R_a^A R_b^A\rangle_{AA} \\
&- \langle x_{1a}x_{2b}\rangle_{AA} - \langle x_{1a}R_b^A\rangle_{AA} - \langle R_a^A x_{2b}\rangle_{AA} - \langle R_a^A R_b^A\rangle_{AA} \qquad \text{(A–25)} \\
&= \langle x_{1a}x_{1b}\rangle_{AA} - \langle x_{1a}x_{2b}\rangle_{AA} + R_a^A\langle x_{1b}\rangle_{AA} - R_a^A\langle x_{2b}\rangle_{AA} \qquad \text{(A–26)}
\end{aligned}$$

The last two contributions are equal and of opposite sign, so that

$$\Lambda_{ab}^{AA} = \langle x_{1a}x_{1b}\rangle_{AA} - \langle x_{1a}x_{2b}\rangle_{AA} \qquad \text{(A–27)}$$

1 In what follows the tildes (˜) will be suppressed for simplicity with the coordinates being always local, see figure 56

A.3 BASIC INTEGRALS IN TERMS OF MONADIC FUNCTIONS. IMPLEMENTATION IN `promolden`

In this section we will evaluate the general integral

$$\langle x_{1a}^{n} x_{1a'}^{n'} x_{2b}^{m} x_{2b'}^{m'} \rangle_{AB} = \int_{A} d\mathbf{r}_1 x_{1a}^{n} x_{1a'}^{n'} \int_{B} d\mathbf{r}_2 x_{2b}^{m} x_{2b'}^{m'} \rho_{xc}(\mathbf{r}_1, \mathbf{r}_2), \tag{A–28}$$

with all possible combinations of the exponents n, n', m, and m'. For this purpose, we express $\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2)$ in terms of the monadic functions [146]

$$\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i,j}^{M} \eta_{ij} f_{ij}(\mathbf{r}_1) f_{ij}(\mathbf{r}_2), \tag{A–29}$$

where M is the number of partially of fully occupied MOs, $\eta_{ij} = \eta_{ji}$ are known coefficients, and $f_{ij} = f_{ji}$ is a known linear combination of producs of MOs φ_i. For closed-shell 1-det molecules, $\eta_{ij} = -2$, $f_{ij} = \varphi_i \varphi_j$, and $M = N/2$, where N is the number of electrons. Then

$$\langle x_{1a}^{n} x_{1a'}^{n'} x_{2b}^{m} x_{2b'}^{m'} \rangle_{AB} = \sum_{i,j} \eta_{ij} \int_{A} d\mathbf{r}_1 f_{ij}(\mathbf{r}_1) x_{1a}^{n} x_{1a'}^{n'} \int_{B} d\mathbf{r}_2 f_{ij}(\mathbf{r}_2) x_{2b}^{m} x_{2b'}^{m'}, \tag{A–30}$$

$$\equiv \sum_{i,j} \eta_{ij} \int_{A} d\mathbf{r} f_{ij}(\mathbf{r}) x_{a}^{n} x_{a'}^{n'} \int_{B} d\mathbf{r} f_{ij}(\mathbf{r}) x_{2}^{m} x_{b'}^{m'} \tag{A–31}$$

$$\equiv \sum_{i,j} \eta_{ij} \langle x_{a}^{n} x_{a'}^{n'} \rangle_{ij,A} \langle x_{b}^{m} x_{b'}^{m'} \rangle_{ij,B} \tag{A–32}$$

According to the expressions in Section A.2 the following basic integrals over every domain Ω are necessary:

$$\langle\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) = \int r^2 dr \int_{\hat{r}} f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–33}$$

$$\langle x\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) x = \int r^2 dr \int_{\hat{r}} x f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–34}$$

$$\langle y\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) y = \int r^2 dr \int_{\hat{r}} y f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–35}$$

$$\langle z\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) z = \int r^2 dr \int_{\hat{r}} z f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–36}$$

$$\langle x^2\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) x^2 = \int r^2 dr \int_{\hat{r}} x^2 f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–37}$$

$$\langle y^2\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) y^2 = \int r^2 dr \int_{\hat{r}} y^2 f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–38}$$

$$\langle z^2\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) z^2 = \int r^2 dr \int_{\hat{r}} z^2 f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–39}$$

$$\langle xy\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) xy = \int r^2 dr \int_{\hat{r}} xy f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–40}$$

$$\langle xz\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) xz = \int r^2 dr \int_{\hat{r}} xz f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r} \tag{A–41}$$

$$\langle yz\rangle_{ij,\Omega} = \int_\Omega d\mathbf{r} f_{ij}(\mathbf{r}) yz = \int r^2 dr \int_{\hat{r}} yz f_{ij}^{\Omega}(r,\theta,\phi) d\hat{r}, \tag{A–42}$$

$$\tag{A–43}$$

where $d\hat{r} = \sin\theta d\theta d\phi$. In case of a 1-det wavefunction, the first integral $\langle\rangle_{ij,\Omega}$ is equal to the atomic overlap matrix (AOM) element $S_{ij}^{\Omega} = \langle i|j\rangle_{\Omega}$.

We will see now how the above integrals can be obtained from the currently stored integrals in the `promolden` code. There, the following angular averaged are stored:

$$R_{lm}^{\Omega}(r) = \left(\frac{4\pi}{2l+1}\right)^{\frac{1}{2}} \int_{\hat{r}} S_{lm}(\hat{r}) f^{\Omega}(\mathbf{r}) d\hat{r}, \tag{A–44}$$

where

$$f^{\Omega}(\mathbf{r}) = \begin{cases} f(\mathbf{r}) & \text{for } \mathbf{r} \in \Omega \\ 0 & \text{for } \mathbf{r} \notin \Omega. \end{cases} \tag{A–45}$$

and S_{lm} are the real spherical harmonics, defined according to reference [20]. The explicit $R^{\Omega}_{lm}(r)$'s for $l \leqslant 2$ and $-l \leqslant m \leqslant +l$ are

$$R^{\Omega}_{00}(r) = \int_{\hat{r}} f^{\Omega}(\mathbf{r})\,d\hat{r} = \int_{\hat{r}} \frac{x^2+y^2+z^2}{r^2} f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–46}$$

$$R^{\Omega}_{1-1}(r) = \int_{\hat{r}} \frac{y}{r} f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–47}$$

$$R^{\Omega}_{10}(r) = \int_{\hat{r}} \frac{z}{r} f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–48}$$

$$R^{\Omega}_{11}(r) = \int_{\hat{r}} \frac{x}{r} f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–49}$$

$$R^{\Omega}_{2-2}(r) = \int_{\hat{r}} \frac{\sqrt{3}xy}{r^2} f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–50}$$

$$R^{\Omega}_{2-1}(r) = \int_{\hat{r}} \frac{\sqrt{3}yz}{r^2} f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–51}$$

$$R^{\Omega}_{20}(r) = \int_{\hat{r}} \frac{1}{2}\left[\frac{3z^2}{r^2}-1\right] f^{\Omega}(\mathbf{r})\,d\hat{r} = \int_{\hat{r}} \frac{1}{2}\left[\frac{2z^2-x^2-y^2}{r^2}\right] f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–52}$$

$$R^{\Omega}_{21}(r) = \int_{\hat{r}} \frac{\sqrt{3}xz}{r^2} f^{\Omega}(\mathbf{r})\,d\hat{r}, \tag{A–53}$$

$$R^{\Omega}_{22}(r) = \int_{\hat{r}} \frac{\sqrt{3}}{2}\left[\frac{x^2-y^2}{r^2}\right] f^{\Omega}(\mathbf{r})\,d\hat{r}. \tag{A–54}$$

From (A–46)-(A–49) and (A–50), (A–53), (A–51) we have, respectively

$$\langle\rangle_{\Omega} = \int r^2\,dr \times R^{\Omega}_{00}(r) \tag{A–55}$$

$$\langle x\rangle_{\Omega} = \int r^2\,dr \times r R^{\Omega}_{11}(r) \tag{A–56}$$

$$\langle y\rangle_{\Omega} = \int r^2\,dr \times r R^{\Omega}_{1-1}(r) \tag{A–57}$$

$$\langle z\rangle_{\Omega} = \int r^2\,dr \times r R^{\Omega}_{10}(r), \tag{A–58}$$

and

$$\langle xy\rangle_{\Omega} = \int r^2\,dr \times \frac{r^2}{\sqrt{3}} R^{\Omega}_{2-2}(r) \tag{A–59}$$

$$\langle xz\rangle_{\Omega} = \int r^2\,dr \times \frac{r^2}{\sqrt{3}} R^{\Omega}_{21}(r) \tag{A–60}$$

$$\langle yz\rangle_{\Omega} = \int r^2\,dr \times \frac{r^2}{\sqrt{3}} R^{\Omega}_{2-1}(r). \tag{A–61}$$

$$\tag{A–62}$$

From the sum $R^{\Omega}_{00}(r) + 2R^{\Omega}_{20}(r)$, we obtain

$$\int_{\hat{r}} z^2 f^{\Omega}(\mathbf{r})\,d\hat{r} = \frac{r^2}{3}\left[R^{\Omega}_{00}(r) + 2R^{\Omega}_{20}(r)\right]. \tag{A–63}$$

On the other hand, we have

$$R^{\Omega}_{00}(r) + \frac{2}{\sqrt{3}} R^{\Omega}_{22}(r) = \int_{\hat{r}} \frac{2x^2 + z^2}{r^2} f^{\Omega}(\mathbf{r}) d\hat{r} = 2\int_{\hat{r}} \frac{x^2}{r^2} f^{\Omega}(\mathbf{r}) d\hat{r} + \int_{\hat{r}} \frac{z^2}{r^2} f^{\Omega}(\mathbf{r}) d\hat{r}. \tag{A–64}$$

From the above equation and (A–63) we have

$$\int_{\hat{r}} x^2 f^{\Omega}(\mathbf{r}) d\hat{r} = \frac{r^2}{3}\left[R^{\Omega}_{00}(r) + \sqrt{3} R^{\Omega}_{22}(r) - R^{\Omega}_{20}(r)\right]. \tag{A–65}$$

Finally, from (A–54)

$$\int_{\hat{r}} \frac{y^2}{r^2} f^{\Omega}(\mathbf{r}) d\hat{r} = \int_{\hat{r}} \frac{x^2}{r^2} f^{\Omega}(\mathbf{r}) d\hat{r} - \frac{2}{\sqrt{3}} R^{\Omega}_{22}(r), \tag{A–66}$$

and using (A–65) we obtain

$$\int_{\hat{r}} y^2 f^{\Omega}(\mathbf{r}) d\hat{r} = \frac{r^2}{3}\left[R^{\Omega}_{00}(r) - \sqrt{3} R^{\Omega}_{22}(r) - R^{\Omega}_{20}(r)\right]. \tag{A–67}$$

In terms of the above radial integrals the ij contribution of all the ab components of Λ^{AB}_{ab} and Λ^{AA}_{ab} are given by

$$\begin{aligned}
\Lambda^{AB}_{xx} &= \langle x^2\rangle_A \langle\rangle_B + \langle x^2\rangle_B \langle\rangle_A - 2\langle x\rangle_A \langle x\rangle_B - 2R^{AB}_x [\langle x\rangle_A \langle\rangle_B - \langle\rangle_A \langle x\rangle_B] - R^{AB}_x R^{AB}_x N_{AB} \\
\Lambda^{AB}_{yy} &= \langle y^2\rangle_A \langle\rangle_B + \langle y^2\rangle_B \langle\rangle_A - 2\langle y\rangle_A \langle y\rangle_B - 2R^{AB}_y [\langle y\rangle_A \langle\rangle_B - \langle\rangle_A \langle y\rangle_B] - R^{AB}_y R^{AB}_y N_{AB} \\
\Lambda^{AB}_{zz} &= \langle z^2\rangle_A \langle\rangle_B + \langle z^2\rangle_B \langle\rangle_A - 2\langle z\rangle_A \langle z\rangle_B - 2R^{AB}_z [\langle z\rangle_A \langle\rangle_B - \langle\rangle_A \langle z\rangle_B] - R^{AB}_z R^{AB}_z N_{AB} \\
\Lambda^{AB}_{xy} &= \langle xy\rangle_A \langle\rangle_B + \langle xy\rangle_B \langle\rangle_A - \langle x\rangle_A \langle y\rangle_B - \langle x\rangle_B \langle y\rangle_A - R^{AB}_y [\langle x\rangle_A \langle\rangle_B - \langle\rangle_A \langle x\rangle_B] \\
&\quad - R^{AB}_x [\langle y\rangle_A \langle\rangle_B - \langle\rangle_A \langle y\rangle_B] + R^{AB}_x R^{AB}_y N_{AB} = \Lambda^{AB}_{yx} \\
\Lambda^{AB}_{xz} &= \langle xz\rangle_A \langle\rangle_B + \langle xz\rangle_B \langle\rangle_A - \langle x\rangle_A \langle z\rangle_B - \langle x\rangle_B \langle z\rangle_A - R^{AB}_z [\langle x\rangle_A \langle\rangle_B - \langle\rangle_A \langle x\rangle_B] \\
&\quad - R^{AB}_x [\langle z\rangle_A \langle\rangle_B - \langle\rangle_A \langle z\rangle_B] + R^{AB}_x R^{AB}_z N_{AB} = \Lambda^{AB}_{zx} \\
\Lambda^{AB}_{yz} &= \langle yz\rangle_A \langle\rangle_B + \langle yz\rangle_B \langle\rangle_A - \langle y\rangle_A \langle z\rangle_B - \langle y\rangle_B \langle z\rangle_A - R^{AB}_z [\langle y\rangle_A \langle\rangle_B - \langle\rangle_A \langle y\rangle_B] \\
&\quad - R^{AB}_y [\langle z\rangle_A \langle\rangle_B - \langle\rangle_A \langle z\rangle_B] + R^{AB}_y R^{AB}_z N_{AB} = \Lambda^{AB}_{zy}
\end{aligned} \tag{A–68}$$

$$\begin{aligned}
\Lambda^{AA}_{xx} &= \langle x^2\rangle_A \langle\rangle_A - \langle x\rangle^2_A \\
\Lambda^{AA}_{yy} &= \langle y^2\rangle_A \langle\rangle_A - \langle y\rangle^2_A \\
\Lambda^{AA}_{zz} &= \langle z^2\rangle_A \langle\rangle_A - \langle z\rangle^2_A \\
\Lambda^{AA}_{xy} &= \langle xy\rangle_A \langle\rangle_A - \langle x\rangle_A \langle y\rangle_A = \Lambda^{AA}_{yx} \\
\Lambda^{AA}_{xz} &= \langle xz\rangle_A \langle\rangle_A - \langle x\rangle_A \langle z\rangle_A = \Lambda^{AA}_{zx} \\
\Lambda^{AA}_{yz} &= \langle yz\rangle_A \langle\rangle_A - \langle y\rangle_A \langle z\rangle_A = \Lambda^{AA}_{zy}.
\end{aligned} \tag{A–69}$$

A.4 MULTIPOLAR EXPANSION OF EXCHANGE INTERACTIONS

We derive in this appendix equation (5–15), the multipolar approximation to the exact exchange-correlation interaction, equations (5–10) and (5–11). Further details are given in I. We start by using the bipolar expansion for r_{12}^{-1},

$$r_{12}^{-1} = \sum_{l_1 m_1}^{\infty} \sum_{l_2 m_2}^{\infty} S_{l_1 m_1}(\hat{r}_1)\, S_{l_2 m_2}(\hat{r}_2)\, D_{l_1 m_1}^{l_2 m_2}(r_1, r_2, \mathbf{R}), \tag{A–70}$$

where $\mathbf{r}_1 \equiv (r_1, \hat{r}_1)$ and $\mathbf{r}_2 \equiv (r_2, \hat{r}_2)$ are referred to centers A and B, respectively, $\mathbf{R} = (\mathbf{R}_B - \mathbf{R}_A) \equiv (R, \hat{R})$ is the position vector of center B with respect to center A (see figure 56), $S_{lm}(\hat{r})$ are real spherical harmonics defined as[34]

$$S_{lm}(\theta, \phi) = \Theta_{l|m|}(\theta)\Phi_m(\phi), \tag{A–71}$$

$$\Theta_{lm}(\theta) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}\, P_l^m(\cos\theta), \tag{A–72}$$

$$\Phi_m(\phi) = \begin{cases} \sqrt{2}\cos m\phi & m > 0, \\ 1 & m = 0, \\ \sqrt{2}\sin|m|\phi & m < 0, \end{cases} \tag{A–73}$$

and $P_l^m(\cos\theta)$ are the associated Legendre functions, defined for $m \geqslant 0$ by

$$P_l^m(x) = \frac{1}{2^l l!}(1-x^2)^{m/2}\frac{d^{\,l+m}}{dx^{\,l+m}}(x^2-1)^l. \tag{A–74}$$

Finally, $D_{l_1 m_1}^{l_2 m_2}(r_1, r_2, \mathbf{R})$ in equation A–70 is defined as

$$D_{l_1 m_1}^{l_2 m_2}(r_1, r_2, \mathbf{R}) = 4\pi(-1)^{l_1} \sum_{l_3 = |l_1 - l_2|}^{l_1 + l_2} \mathcal{V}_{l_1 l_2 l_3}(r_1, r_2, R)\, T_{l_1 m_1 l_2 m_2}^{l_3}(\hat{R}), \tag{A–75}$$

where the sum over l_3 runs in steps of 2, $\mathcal{V}_{l_1 l_2 l_3}(r_1, r_2, R)$ is a discriminant that takes different expressions in the four regions defined in figure 57, and $T_{l_1 m_1 l_2 m_2}^{l_3}(\hat{R})$ is the angular factor

$$T_{l_1 m_1 l_2 m_2}^{l_3}(\hat{R}) = \sum_{m_3 = -l_3}^{+l_3} d_{l_1 m_1 l_2 m_2}^{l_3 m_3}\, S_{l_3 m_3}(\hat{R}), \tag{A–76}$$

where $d_{l_1 m_1 l_2 m_2}^{l_3 m_3}$ is the Gaunt coefficient between the $S_{lm}(\theta, \phi)$'s defined by

$$d_{l_1 m_1 l_2 m_2}^{l_3 m_3} = \langle S_{l_3 m_3} | \, S_{l_1 m_1} \, | S_{l_2 m_2} \rangle. \tag{A–77}$$

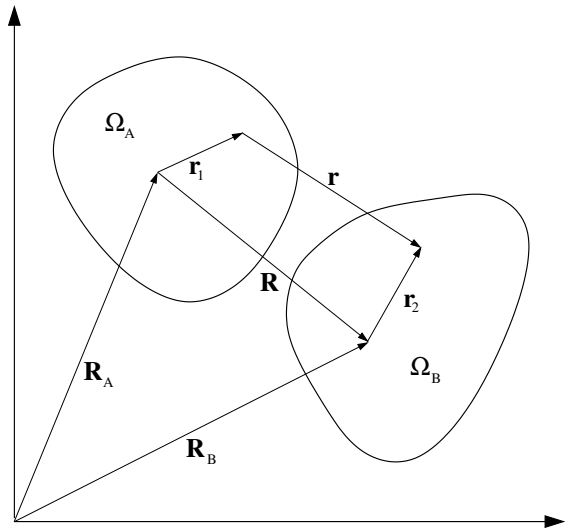


Figure 56: Coordinate System

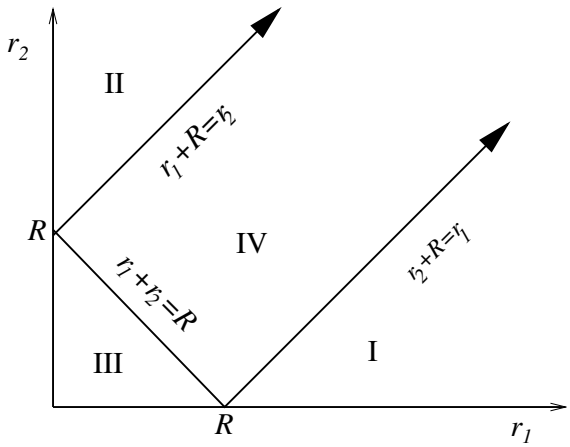


Figure 57: Regions of definition of the $\mathcal{V}_{l_1 l_2 l_3}(r_1, r_2, R)$ function

Given that S_{lm} is real, $d^{l_3 m_3}_{l_1 m_1 l_2 m_2}$ is invariant against any permutation of the pair of indices (l_i, m_i). These coefficients may be determined as described elsewhere. Using (A–70) in (5–11) one has

$$K^{AB}_{ij} = \sum_{l_1 m_1} \sum_{l_2 m_2} \int_{\Omega_A} S_{l_1 m_1}(\hat{r}_1)\, f_{ij}(\mathbf{r}_1)\, d\mathbf{r}_1 \int_{\Omega_B} S_{l_2 m_2}(\hat{r}_2)\, f_{ij}(\mathbf{r}_2)\, d\mathbf{r}_2\, D^{l_2 m_2}_{l_1 m_1}(r_1, r_2, \mathbf{R})). \tag{A–78}$$

A further simplification of K^{AB}_{ij} requires the explicit form of $D^{l_2 m_2}_{l_1 m_1}(r_1, r_2, \mathbf{R})$. From the expression of $\mathcal{V}_{l_1 l_2 l_3}(r_1, r_2, R)$ (equations (B1)-(B9) of I) it follows that, as long as $r_1 + r_2 \leqslant R$, this discriminant only takes a nonzero value in region III of the (r_1, r_2) space (see figure 57). This condition will be exactly satisfied if $R \geqslant r_1^{max} + r_2^{max}$, where r_1^{max} is the maximum value of the radial coordinate within Ω_A, with an equivalent definition for r_2^{max}. In the present context, atoms A and B are said to be non-overlapping if this condition is fulfilled, and overlapping otherwise. Although it may occur that the condition $R \geqslant r_1^{max} + r_2^{max}$ is not exactly satisfied, provided that the atomic basins Ω_A and Ω_B are well-separated in the space, we can expect that it is fulfilled in practical terms. The multipolar approach, intensively used to approximate the Coulomb repulsion in the modellization of biomolecules, is equivalent to the assumption that $r_1 + r_2 \leqslant R$ for any r_1 and r_2. Thus, region III is identified with the complete first quadrant. In this region, $D^{l_2 m_2}_{l_1 m_1}(r_1, r_2, \mathbf{R})$ is given by

$$D^{l_2 m_2}_{l_1 m_1}(r_1, r_2, \mathbf{R}) = (-1)^{l_1} 16\pi^2\, \Delta_{l_1 l_2}\, \frac{r_1^{l_1} r_2^{l_2}}{R^{l_1+l_2+1}}\, T^{l_1+l_2}_{l_1 m_1 l_2 m_2}(\hat{R}), \tag{A–79}$$

where

$$\Delta_{l_1 l_2} = (-1)^{l_1+l_2} \frac{(2l_1 + 2l_2)!\, l_1!\, l_2!}{(l_1 + l_2)!\, (2l_1 + 1)!\, (2l_2 + 1)!}. \tag{A–80}$$

Using (A–79) in (A–78) we get

$$\left(K^{AB}_{ij}\right)_{lr} = \sum_{l_1 m_1} \sum_{l_2 m_2} C_{l_1 m_1 l_2 m_2}(\hat{R}) \frac{q^{\Omega_A}_{ij,l_1 m_1}\, q^{\Omega_B}_{ij,l_2 m_2}}{R^{l_1+l_2+1}}, \tag{A–81}$$

where

$$C_{l_1 m_1, l_2 m_2}(\hat{R}) = (-1)^{l_1} 4\pi\, [(2l_1 + 1)(2l_2 + 1)]^{\frac{1}{2}}\, \Delta_{l_1 l_2}\, T^{l_1+l_2}_{l_1 m_1 l_2 m_2}(\hat{R}), \tag{A–82}$$

and the $q^{\Omega}_{ij,lm}$ have been defined in (5–17). Finally, substituting (A–81) in (5–10) we obtain (5–15), the multipolar approximation for the exchange-correlation interaction, $\left(V^{AB}_{xc}\right)_{lr}$.

B | LIST OF PUBLICATIONS

A Multipolar Approach to the Interatomic Covalent Interaction Energy

Evelio Francisco, Daniel Menéndez Crespo, Aurora Costales, and Ángel Martín Pendás

Interatomic exchange-correlation energies correspond to the covalent energetic contributions to an interatomic interaction in real space theories of the chemical bond, but their widespread use is severely limited due to their computationally intensive character. In the same way as the multipolar (*mp*) expansion is customary used in biomolecular modeling to approximate the classical Coulomb interaction between two charge densities $\rho_A(\mathbf{r})$ and $\rho_B(\mathbf{r})$, we examine in this work the *mp* approach to approximate the interatomic exchange-correlation (xc) energies of the Interacting Quantum Atoms method. We show that the full *xc mp* series is quickly divergent for directly bonded atoms (1–2 pairs) albeit it works reasonably well most times for 1– *n* ($n > 2$) interactions. As with conventional perturbation theory, we show numerically that the *xc* series is asymptotically convergent and that, a truncated *xc mp* approximation retaining terms up to $l_1 + l_2 = 2$ usually gives relatively accurate results, sometimes even for directly bonded atoms. Our findings are supported by extensive numerical analyses on a variety of systems that range from several standard hydrogen bonded dimers to typically covalent or aromatic molecules. The exact algebraic relationship between the monopole-monopole *xc mp* term and the inter-atomic bond order, as measured by the delocalization index of the quantum theory of atoms in molecules, is also established. © 2017 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24758

Introduction

The role of the quantum mechanical exchange-correlation (*xc*) energy as the basic glue binding together atoms and molecules has been clearly stressed in the past.[1] In the chemical literature, however, this insight is less well-known. Although exchange-correlation functionals, for instance, are the essential ingredients in modern implementations of Density Functional Theory (DFT),[2] not much work has been devoted to examine the importance of the exchange-correlation energy itself in the theory of chemical bonding from the DFT viewpoint.[3]

Actually, almost all that is known about the chemical relevance of the *xc* energy has been derived in the last decade through the study of bonding in real or position space.[4,5] With this term, we gather together a number of techniques that are being actively explored[6–8] which use orbital invariant reduced densities (or density matrices) to develop a new paradigm that may one day replace the standard molecular orbital (MO) approach.[9] Usually, these techniques use a partition of real space into regions endowed with chemical meaning, be them atoms, bonds, cores, lone pairs, and so forth. In many cases, the space is divided using the topology induced by the gradient field of an orbital invariant scalar, like the electron density (which gives rise to the atomic partioning of the quantum theory of atoms in molecules (QTAIM) developed by Bader,[10] or the electron localization function (that isolates core, bond and lone pair regions).[11,12] When this topological tools are used, we say that we are under the Quantum Chemical Topology umbrella.[13]

In the context of the QTAIM/QCT, we proposed a number of years ago an exact, general decomposition of the total molecular energy *E* into atomic and inter-atomic terms that we called the interacting quantum atoms (IQA) approach.[4,5] All the expectation values of the standard Coulomb Hamiltonian that make up *E* are written in IQA as a sum of domain contributions, and *E* is obtained by adding atomic self-energies, which tend to the free atomic energies when the atoms that interact are sufficiently far apart, and pairwise additive interaction energies. The latter are composed of a classical term that depends only on classical electrostatic contributions, and an exchange-correlation energy, V_{xc} which accounts for the quantum mechanical effects. As we and others have shown over the years,[14,15] the classical part of the interaction measures its ionic component, while the V_{xc} energy is to be associated with its covalent counterpart.

In these years, the interatomic *xc* energy has become an important ingredient of any quantitative account of chemical bonding in position space.[16,17] For instance, it has been shown to be intimately related to the appearance of the bond critical points of the QTAIM, leading to the concept of priviledged exchange-correlation channels.[18] It has also been used to reconstruct molecular graphs from purely energetic quantities,[19] to shed light on new concepts like halogen bonding,[20,21] to recover stereolectronic effects,[22] or to find new long-range electronic anomalies.[23]

Interatomic V_{xc} energies are intimately linked to the delocalization or shared electron delocalization indices (DIs) used in

Evelio Francisco, Daniel Menéndez Crespo, Aurora Costales, Ángel Martín Pendás
Departamento de Química Física y Analítica, Facultad de Química. Universidad de Oviedo, Oviedo 33006, Spain
Contract grant sponsor: Spanish MINECO; Contract grant number: CTQ2015-65790-P; Contract grant sponsor: FICyT; Contract grant number: GRUPIN14-049; Contract grant sponsor: European Union FEDER funds

the QTAIM, defined almost 40 years ago by Bader and Stephens.[24] These are obtained by directly integrating the *xc* density of very two different atomic domains and measure the number of shared pairs of electrons between them. They have been successfully used as real space generalization of the bond order concept, reducing to the Wiberg-Mayer[25,26] bond orders if atomic domains are imagined to collapse onto their nuclei. In a sense, V_{xc}'s are the energetic counterparts of DIs, and both have been empirically found to correlate very well when a given couple of atoms is examined in different molecular environments.

The computational complexity of obtaining DIs is considerably smaller than that of calculating V_{xc}'s, as the former may be factorized into sums of products of atomic overlap matrices (AOM; 3D numerical integrals), while the latter need, in principle, very costly 6D quadratures. Thus, if we are not interested in very accurate results, but only in semi-quantitative estimations of covalent energies, any procedure that might approximate the V_{xc} values in terms of cheaper to compute quantities like the DIs should be welcome. That procedure was initially examined by Rafat and Popelier,[27] that wrote each interatomic Hartree–Fock (HF) V_{xc} interaction as a multipolar series, exploring the convergence of this series in different closed-shell molecules computed at this level of theory. In this work, we generalize their algebraic formalism to multi-determinant wavefunctions. This generalization is possible thanks to the use of the monadic diagonalization of the exchange-correlation density,[28] customary used within the IQA methodology.[4,5] Our expressions converge to those of Ref. [27] when HF exchange-correlation densities are used in the calculation. We will show that, regardless the type of calculation, the monopole-monopole term of the multipolar *xc* interaction between two atoms A and B of the molecule coincides with that of Rafat and Popelier, being equal to $-\delta^{AB}/(2R)$ (R is the AB internuclear distance), and that the series is usually divergent although many times asymptotically convergent. Moreover, our results clearly establish in what conditions V_{xc} can be safely approximated by a truncated series, and how in some situations retaining up to the charge-quadrupole terms may give reasonable results even for directly bonded atoms. In the latter cases, the use of the crudest approach $V_{xc}^{AB} \sim -\delta^{AB}/(2R)$ is even preferable to using the multipolar expansion up to very high order.

We will first consider the multipolar expansion of V_{xc}, including a short account of the IQA methodology. Then, we will turn to examine how the series converges or diverges for a number of selected systems.

Multipolar Expansion of V_{xc}^{AB}

In this section, we briefly describe the IQA method and the role played by the exchange–correlation (*xc*) interaction in this energy partition method (Subsection), the exact computation of this interaction (Subsection), and its multipolar approximation (MP) with or without truncating the expansion of the angular momentum series (Subsection). It is worth noting that the experience gained to date with the IQA method, both by us and by other groups, clearly indicates that the magnitude of V_{xc}^{AB} correlates very well with the degree of covalency between the pair of atoms A and B as measured by means of the DI defined by Bader and Stephens, and weighted through the inverse of the distance between both atoms. As we will see, this correlation would be perfect as long as the crudest MP to V_{xc}^{AB} (consisting in truncating the multipolar series in the term $l_1=m_1=l_2=m_2=0$) were exact.

The IQA method

The IQA method[4,5] is a real space energetic partition inspired in the QTAIM that focuses on domain-averaged integrated quantities. The total energy in this approach is given by

$$E=\sum_A T_A+V_{en}^{AA}+V_{ee}^{AA}+\sum_{A>B} V_{nn}^{AB}+V_{en}^{AB}+V_{en}^{BA}+V_{ee}^{AB} \quad (1)$$

$$=\sum_A E_{self}^{A}+\sum_{A>B} E_{int}^{AB}. \quad (2)$$

where A runs over all the atoms in the molecule, $V_{nn}^{AB}=Z^A Z^B/R_{AB}$ is the repulsion between the nuclei A and B, $V_{en}^{AB}=-Z^B\int_{\Omega_A} d\boldsymbol{r}_1\,\rho(\boldsymbol{r}_1)\boldsymbol{r}_{1B}^{-1}$ is the nuclear attraction of the electrons within the basin of A (Ω_A) to the nucleus B, and V_{ee}^{AB} is the total electron repulsion between Ω_A and Ω_B. The latter is given by $V_{ee}^{AB}=J^{AB}+V_{xc}^{AB}$ where

$$J^{AB}=\int_{\Omega_A} d\mathbf{r}_1\int_{\Omega_B} d\mathbf{r}_2\, r_{12}^{-1}\,\rho(\mathbf{r}_1)\,\rho(\mathbf{r}_2), \quad (3)$$

is the classical or Coulomb electron-electrons repulsion, and

$$V_{xc}^{AB}=\int_{\Omega_A} d\mathbf{r}_1\int_{\Omega_B} d\mathbf{r}_2\, r_{12}^{-1}\,\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2), \quad (4)$$

where $\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2)$ is the exchange-correlation (*xc*) density, is the purely quantum-mechanical electron-electron *xc* interaction, which is the main subject of this work. In this way,

$$E_{int}^{AB}=\left(V_{nn}^{AB}+V_{en}^{AB}+V_{en}^{BA}+J^{AB}\right)+V_{xc}^{AB}=V_{cl}^{AB}+V_{xc}^{AB}. \quad (5)$$

The term E_{self}^{A} in eq. (2) collects all the energetic components affecting exclusively to the atom A while E_{int}^{AB} represents the full interaction energy between atoms A and B, that is made of the full electrostatic or classical interaction (V_{cl}^{AB}) and the quantum-mechanical part (V_{xc}^{AB}). The expression 2 is valid, not only for the IQA methodology but also for other energetic partitions, such as a recently proposed one inspired in the IQA method, although using a fuzzy partition of the space and localized MO.[29] Mayer and Hamza have also dealt with the exchange component in eq. (4) in the framework of a Hilbert space partition instead of the real space QTAIM partition we use here.[30]

The exact *xc* interaction energy

Over the years, it has become clear that the magnitude of V_{xc}^{AB} measures the degree of covalency of the chemical bond between the atoms A and B. The more negative its value, the

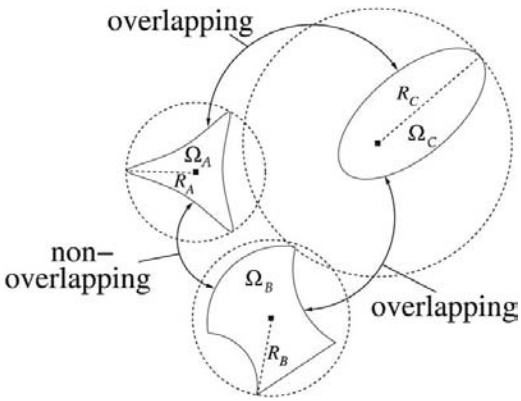


Figure 1. Schematic representation of overlapping and non-overlapping regions.

bigger the bond order between the two atoms and vice versa.[6,14,18] Their values have been recently proposed as a novel solution to the problem of assigning a molecular graph to a collection of nuclei[23] (i.e., how to draw a molecular structure). In the IQA approach, this term is exactly computed as follows. First, we use the fact that for both single- (1-det) and multi-determinant wavefunctions built in with real MOs ϕ_i, $\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)$ can be written as

$$\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2)=\sum_{i,j,k,l}^{M}\lambda_{ijkl}\,\phi_i(\boldsymbol{r}_1)\,\phi_j(\boldsymbol{r}_1)\,\phi_k(\boldsymbol{r}_2)\,\phi_l(\boldsymbol{r}_2), \tag{6}$$

where M is the number of partially or fully occupied MOs, and λ_{ijkl} is a symmetric matrix in the (i, j) and (k, l) pairs. Defining a set of coefficients,

$$\varepsilon_{ijkl}=\lambda_{ijkl}+\lambda_{jikl}(1-\delta_{ij})+\lambda_{ijlk}(1-\delta_{kl})+\lambda_{jilk}(1-\delta_{ij})(1-\delta_{kl}), \tag{7}$$

where δ_{ij} is the Kronecker symbol ($\delta_{ij}=1$ for $i=j$, $\delta_{ij}=0$ for $i\neq j$) we may write an (i, j),(k, l) symmetric simpler expression,

$$\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2)=\sum_{i\geq j,k\geq l}^{M}\varepsilon_{ijkl}\phi_i(\boldsymbol{r}_1)\,\phi_j(\boldsymbol{r}_1)\,\phi_k(\boldsymbol{r}_2)\,\phi_l(\boldsymbol{r}_2). \tag{8}$$

Using the basis of products of MOs, $\{\phi_i(\boldsymbol{r})\phi_j(\boldsymbol{r}), i\geq j\}$, that contains $M(M+1)/2$ members, we diagonalize eq. (8), and get[28]:

$$\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)=\sum_{i\geq j}^{M}\eta_{ij}f_{ij}(\mathbf{r}_1)f_{ij}(\mathbf{r}_2), \tag{9}$$

where the f_{ij} eigenfunctions are linear combinations of the above products. The ε matrix may be easily computed from the explicit form of a given calculated wavefunction. For closed-shell 1-det wavefunctions (and formally also for a Kohn-Sham determinant) $M=N/2$, where N is the number of electrons, the ε matrix is already diagonal in the (i, j) and (k, l) pairs, each eigenvector is the product of two MOs, $f_{ij}=\phi_i\phi_j$, and the η_{ij} eigenvalues are simply $\eta_{ii}=-2$ and $\eta_{ij}=-4$ ($i\neq j$). Using eq. (9) in the expression of V_{xc}^{AB} one gets

$$V_{xc}^{AB}=\sum_{i\geq j}^{M}\eta_{ij}K_{ij}^{AB}, \text{ where} \tag{10}$$

$$K_{ij}^{AB}=\int_{\Omega_A}d\mathbf{r}_1\int_{\Omega_B}d\mathbf{r}_2\,r_{12}^{-1}\,f_{ij}(\mathbf{r}_1)f_{ij}(\mathbf{r}_2). \tag{11}$$

The integrals 3 and 11 can be computed numerically and (in principle) exactly, that is, without invoking any approximation such as the multipolar expansion, by means of the bipolar expansion as described in Ref. [31]. Notice that using the Fock-Dirac exchange from Kohn-Sham determinants is an approximation that has no rigorous justification.

The multipolar approach for V_{xc}^{AB}

Comparing eq. (11) with eq. (3) for the Coulomb repulsion it is evident that if J^{AB} is approximated making use of physically reasonable arguments that are also valid for K_{ij}^{AB}, the steps to approximate the latter will be the same used for J^{AB}. The long-range or MP to J^{AB}, given by

$$J_{lr}^{AB}=\sum_{l_1m_1}^{\infty}\sum_{l_2m_2}^{\infty}C_{l_1m_1l_2m_2}(\hat{R})\,\frac{Q_{l_1m_1}^{\Omega_A}Q_{l_2m_2}^{\Omega_B}}{R^{l_1+l_2+1}}, \tag{12}$$

where m_1 (m_2) runs from $-l_1$ ($-l_2$) to $+l_1$ ($+l_2$), $\mathbf{R}=(\mathbf{R}_B-\mathbf{R}_A)\equiv(R,\hat{R})$, with $R=|\mathbf{R}_B-\mathbf{R}_A|$ and $\hat{R}\equiv(\theta_B-\theta_A,\phi_B-\phi_A)$ (see Fig. 14) is the position vector of the B center with respect to the A center, $C_{l_1m_1l_2m_2}(\hat{R})$ are known coefficients, Q_{lm}^{Ω} are the spherical atomic multipoles, defined as

$$Q_{lm}^{\Omega}=N_l\int_{\Omega}r^l\,S_{lm}(\hat{r})\,\rho(\mathbf{r})\,d\mathbf{r}, \tag{13}$$

$N_l=\sqrt{4\pi/(2l+1)}$, and $S_{lm}(\hat{r})$ are real spherical harmonics (see Appendix) is exact when the basins Ω_A and Ω_B are non-overlapping (see Fig. 1 and the definition of overlapping and non-overlapping regions below). Equation (12) is the same

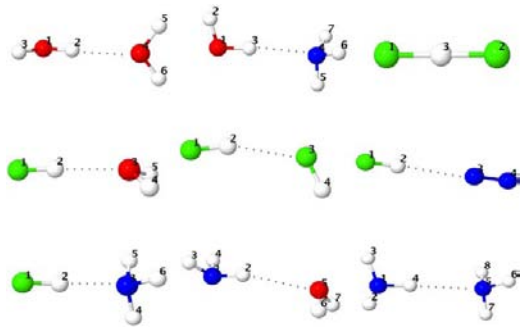


Figure 2. Hydrogen bond systems studied in this work. Hydrogen, nitrogen, oxygen, and fluorine atoms are represented as growing size spheres, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

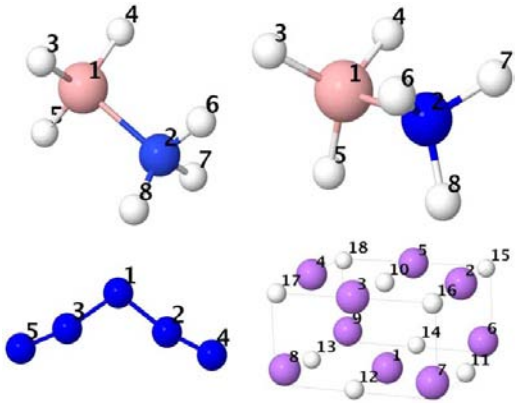


Figure 3. Staggered BH_3NH_3, eclipsed BH_3NH_3, N_5^+, and Li_9H_9 molecules. [Color figure can be viewed at wileyonlinelibrary.com]

used by Popelier et al.[33–35] in their discussion of the multipolar expansion for the diatomic Coulomb repulsion. Retaining only terms with $l_1 \leq 1$ and $l_2 \leq 1$ in this equation one has

$$J^{AB}_{lr,cd} \simeq \frac{Q^A Q^B}{R} - Q^A \frac{\vec{\mu}^B \cdot \mathbf{R}}{R^3} + Q^B \frac{\vec{\mu}^A \cdot \mathbf{R}}{R^3} + \frac{1}{R^3}\left(\vec{\mu}^A \cdot \vec{\mu}^B - 3\frac{(\vec{\mu}^A \cdot \mathbf{R})\,(\mathbf{R} \cdot \vec{\mu}^B)}{R^2}\right), \tag{14}$$

where $Q^\Omega = \int_\Omega \rho(\mathbf{r})d\mathbf{r}$ and $\vec{\mu}^\Omega = \int_\Omega \mathbf{r}\rho(\mathbf{r})d\mathbf{r}$ are the total electron charge and the dipole moment of the Ω region, respectively. The first, second plus third, and fourth terms of 14 correspond to the charge-charge (*cc*), charge-dipole (*cd*), and dipole-dipole (*dd*) interactions, respectively. We should note that the second and third terms have opposite signs.

If the same approximation is used for K^{AB}_{ij}, $\left(V^{AB}_{xc}\right)_{lr}$ becomes

$$\left(V^{AB}_{xc}\right)_{lr} = \sum_{l_1 m_1}^{\infty}\sum_{l_2 m_2}^{\infty} C_{l_1 m_1 l_2 m_2}(\hat{R})\frac{\delta^{AB}_{l_1 m_1, l_2 m_2}}{R^{l_1+l_2+1}}, \quad \text{where} \tag{15}$$

$$\delta^{AB}_{l_1 m_1, l_2 m_2} = \sum_{i\geq j}^{M} \eta_{ij} q^{\Omega_A}_{ij,l_1 m_1} q^{\Omega_B}_{ij,l_2 m_2}, \text{ and} \tag{16}$$

$$q^{\Omega}_{ij,lm} = N_l \int_\Omega r^l S_{lm}(\hat{r}) f_{ij}(\mathbf{r})d\mathbf{r}. \tag{17}$$

It is important to stress that, similarly to $(J^{AB})_{lr}$, the expression 15 for $(V^{AB}_{xc})_{lr}$ provides the exact *xc* interaction when the atomic basins Ω_A and Ω_B do not overlap (Fig. 1). In the present context, these two basins are non-overlapping because the two spheres of radii R_A and R_B, centered at the origin of Ω_A and Ω_B, respectively, do not intersect each other, being R_A (R_B) the maximum distance from the origin of the basin to the surface of Ω_A (Ω_B). On the contrary, Ω_A and Ω_C are overlapping despite that none point inside Ω_A belongs also to Ω_C and vice versa. When the non-overlapping condition is not met, the current expressions for $(J^{AB})_{lr}$ and $(V^{AB}_{xc})_{lr}$ are only conditionally convergent. We will see different examples of this in section.

The function $N_l r^l S_{lm}(\hat{r})$ is 1 for $l=m=0$, (y, z, x) for $l=1$ and $m=(-1,0,+1)$, and $(\sqrt{3}xy, \sqrt{3}yz, \frac{1}{2}(3z^2-r^2), \sqrt{3}xz, \frac{\sqrt{3}}{2}(x^2-y^2))$ for $l=2$ and $m=(-2,-1,0,+1,+2)$. If, as in the case of J^{AB}, only terms with $l_1 \leq 1$ and $l_2 \leq 1$ are included, $(V^{AB}_{xc})_{lr}$ becomes

$$\left(V^{AB}_{xc}\right)_{lr,cd} \simeq \sum_{i,j}\eta_{ij}\left[\frac{q^{\Omega_A}_{ij} q^{\Omega_B}_{ij}}{R} - q^{\Omega_A}_{ij}\frac{\vec{\mu}^{\Omega_B}_{ij}\cdot\mathbf{R}}{R^3} + q^{\Omega_B}_{ij}\frac{\vec{\mu}^{\Omega_A}_{ij}\cdot\mathbf{R}}{R^3} + \frac{1}{R^3}\left(\vec{\mu}^{\Omega_A}_{ij}\cdot\vec{\mu}^{\Omega_B}_{ij} - 3\frac{(\vec{\mu}^{\Omega_A}_{ij}\cdot\mathbf{R})\,(\mathbf{R}\cdot\vec{\mu}^{\Omega_B}_{ij})}{R^2}\right)\right], \tag{18}$$

where $q^{\Omega}_{ij} \equiv q^{\Omega}_{ij,00} = \int_\Omega f_{ij}(\mathbf{r})d\mathbf{r}$, and

$$\vec{\mu}^{\Omega}_{ij} \equiv (q^{\Omega}_{ij,1-1}, q^{\Omega}_{ij,10}, q^{\Omega}_{ij,1+1}) = \int_\Omega \mathbf{r} f_{ij}(\mathbf{r})d\mathbf{r}. \tag{19}$$

If terms with $(l_1=0, l_2=2)$ and $(l_1=2, l_2=0)$ are also included, the extra contribution

$$\left(V^{AB}_{xc}\right)_{lr,cq} = \sum_{i,j}^{M}\frac{\eta_{ij}}{R^3}\sum_{m=-2}^{+2} q_{2m}(\hat{R})\left[q^A_{ij} q^B_{ij,2m} + q^B_{ij} q^A_{ij,2m}\right] \tag{20}$$

must be added to 18. The *cq* subscript in eq. (20) stands for *charge-quadrupole* interactions. The improved expression for $(V^{AB}_{xc})_{lr}$ is then

$$\left(V^{AB}_{xc}\right)_{lr,cdq} = \left(V^{AB}_{xc}\right)_{lr,cd} + \left(V^{AB}_{xc}\right)_{lr,cq}. \tag{21}$$

The physical meaning of q^{Ω}_{ij} and $\vec{\mu}^{\Omega}_{ij}$ are easy to grasp. If we consider the particular case of their diagonal expressions ($i=j$) for a 1-det wavefunction, $f_{ii}(\mathbf{r}) = \phi^2_i(\mathbf{r})$, so that q^{Ω}_{ii} is the electron charge of the orbital distribution $\phi^2_i(\mathbf{r})$ within the Ω region, and $\vec{\mu}^{\Omega}_{ii}$ the dipole moment of Ω due to this distribution. For this reason, q^{Ω}_{ij} and $\vec{\mu}^{\Omega}_{ij}$ may be called orbital overlap charge and orbital overlap dipole, respectively. At the HF level,

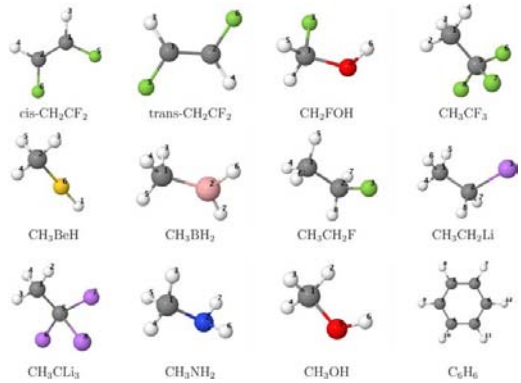


Figure 4. Molecules derived from saturated hydrocarbons by substituting C or H atoms by Be, B, N, O, F atoms, plus the benzene molecule. [Color figure can be viewed at wileyonlinelibrary.com]

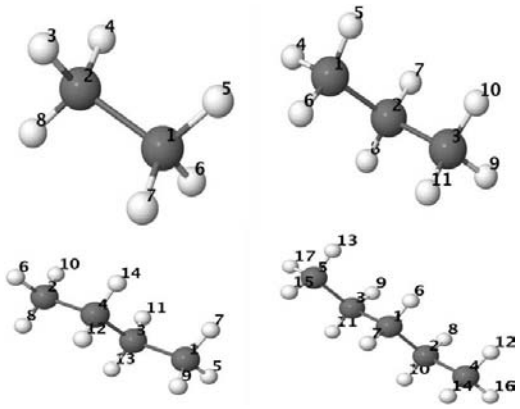


Figure 5. C_nH_{2n+2} (n=2−5) saturated hydrocarbons.

the q_{ij}^{Ω}'s coincide with the AOM elements of the QTAIM, $q_{ij}^{\Omega} \equiv \langle \phi_i | \phi_j \rangle_{\Omega} = S_{ij}^{\Omega}$. However, given that $f_{ij}(\mathbf{r})$ at the correlated level is a linear combination of $\phi_i(\mathbf{r})\phi_j(\mathbf{r})$ products, q_{ij}^{Ω} in this case is a linear combination of AOM elements. Nevertheless, for both types of wavefunctions $-2\delta_{00,00}^{AB}$ coincides with δ^{AB}, the so-called DI between the atoms A and B

$$-2\delta_{00,00}^{AB} = \delta^{AB} = -\sum_{i \geq j}^{M} 2\eta_{ij} S_{ij}^{\Omega_A} S_{ij}^{\Omega_B}, \tag{22}$$

so that the leading term of $(V_{xc}^{AB})_{lr}$ ($l_1 = l_2 = m_1 = m_2 = 0$) can be written as

$$R^{-1} \sum_{i,j} \eta_{ij} q_{ij}^{\Omega_A} q_{ij}^{\Omega_B} = -\delta^{AB}/(2R). \tag{23}$$

The above equation is behind the good existing correlation between the values of V_{xc}^{AB} and δ^{AB} for a large collection of AB couples in many systems. The present derivation shows that the proportionality between V_{xc}^{AB} and δ^{AB} is modulated by the inverse of the distance between the nuclei of both atomic basins.

Systems and Computational Details

All the calculations of this work have performed with our PROMOLDEN code.[36] This program allows the exact computation[28,31] (i.e., without suffering the convergence problems inherent to the multipolar series expansion) of V_{xc}^{AB} as well as the full (lr) and truncated (lr,cd) and (lr,cdq) MP described in section. For brevity, only the exact, and the (lr) and (lr,cdq) numbers will be given in the tables. The errors plotted in the figures are defined as $[(V_{xc}^{AB})_{method} - (V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}| \times 100$, where method=(lr), (lr,cd), or (lr,cdq). The studied systems include several standard hydrogen bonded (HB) dimers (Fig. 2), the staggered BH_3NH_3, eclipsed BH_3NH_3, N_5^+, and Li_9H_9 molecules (Fig. 3), 11 molecules derived from saturated hydrocarbons by substituting C or H atoms by Be, B, N, O, F atoms, plus the benzene molecule (Fig. 4), the saturated hydrocarbons ethane, propane, butane, and pentane (Fig. 5), and the phenol dimer (Fig. 6). The labels of the atoms in the tables are those defined in these figures. For simplicity, the MO required for evaluating the exchange-correlation density of eq. (8) have been obtained through restricted Hartree–Fock (RHF) calculations at the corresponding equilibrium geometries with basis sets of quality 6-311G(d,p) or higher. However, as our results in this article stem from the algebraic properties of the multipolar expansion, we do not expect significant changes neither in the numerical results nor in the subsequent discussion when using more accurate wavefunctions or the approximate data coming from Kohn-Sham determinants in the computation of the *xc* interactions. To prove the validity of this assertion, we will compare the V_x^{AB} energies obtained for staggered ethane in a CAS[14,14] calculation (Complete Active Space calculation with all except the carbon 1s electrons distributed into 14

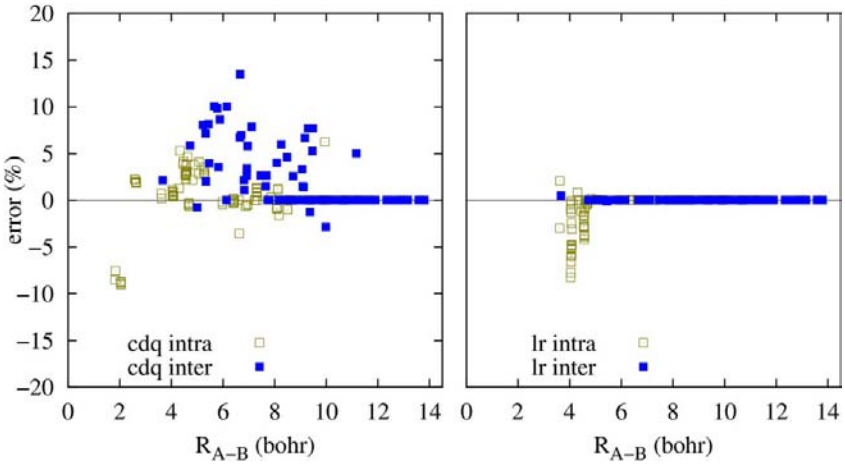


Figure 6. Phenol dimer. [Color figure can be viewed at wileyonlinelibrary.com]

Table 1. *xc* interaction energies $\geq$ 0.1 kJ/mol for the HB dimers of Figure 2.

$A-B$	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	$A-B$	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
		H_2O–H_2O				**HF–NH_3**	
O_1-H_2	−478.59	1.29×10^{4}	−437.77	F_1-H_2	−280.40	3.79×10^{3}	−259.76
O_1-H_3	−569.61	-8.45×10^{5}	−514.46	F_1-N_3	−40.44	−41.74	−41.85
O_1-O_4	−16.40	−16.96	−16.94	F_1-H_4	−0.42	−0.42	−0.42
O_1-H_5	−0.11	−0.10	−0.11	F_1-H_5	−0.42	−0.42	−0.42
H_2-H_3	−1.94	−1.44	−1.84	F_1-H_6	−0.42	−0.42	−0.42
H_2-O_4	−24.73	−25.45	−25.36	H_2-N_3	−47.99	−48.50	−48.23
H_2-H_5	−0.17	−0.17	−0.17	H_2-H_4	−0.34	−0.34	−0.34
H_3-O_4	−0.19	−0.18	−0.18	H_2-H_5	−0.34	−0.34	−0.34
O_4-H_5	−541.14	-7.16×10^{5}	−488.52	H_2-H_6	−0.34	−0.34	−0.34
H_5-H_6	−2.19	−4.27	−2.03	N_3-H_4	−758.46	5.38×10^{5}	−688.30
		H_2O–NH_3		N_3-H_5	−758.46	5.38×10^{5}	−688.30
O_1-H_2	−575.13	-1.21×10^{7}	−519.23	N_3-H_6	−758.40	5.58×10^{5}	−688.24
O_1-H_3	−460.75	-5.35×10^{4}	−421.84	H_4-H_5	−5.34	−12.72	−5.12
O_1-N_4	−20.16	−20.71	−20.70	H_4-H_6	−5.34	−12.71	−5.12
O_1-H_5	−0.19	−0.19	−0.19	H_5-H_6	−5.34	−12.71	−5.12
O_1-H_6	−0.20	−0.20	−0.20			**NH_3-H_2O**	
H_2-H_3	−1.88	−1.00	−1.80	N_1-H_2	−731.84	4.54×10^{6}	−668.43
H_2-N_4	−0.24	−0.24	−0.24	N_1-H_3	−779.59	1.37×10^{6}	−710.26
H_3-N_4	−31.54	−32.23	−32.07	N_1-O_5	−8.33	−8.63	−8.60
H_3-H_5	−0.27	−0.27	−0.27	H_2-H_3	−5.26	−6.18	−5.13
H_3-H_6	−0.28	−0.28	−0.28	H_2-O_5	−16.83	−17.59	−17.37
N_4-H_5	−764.79	6.93×10^{5}	−694.12	H_2-H_6	−0.10	−0.10	−0.10
N_4-H_6	−767.36	5.65×10^{5}	−697.06	H_3-H_4	−6.08	−30.52	−5.85
H_5-H_6	−5.60	−17.52	−5.37	H_3-O_5	−0.20	−0.20	−0.20
H_6-H_7	−5.66	−18.08	−5.43	O_5-H_6	−550.78	-8.08×10^{5}	−496.97
		FHF^-		H_6-H_7	−2.27	−4.67	−2.11
F_1-F_2	−92.79	−100.02	−96.55			**NH_3-NH_3**	
F_1-H_3	−163.82	6.80×10^{4}	−159.19	N_1-H_2	−781.96	7.75×10^{5}	−712.71
		HF–H_2O		N_1-H_4	−719.70	-1.22×10^{5}	−657.38
F_1-H_2	−298.63	−844.89	−275.72	N_1-N_5	−9.02	−9.29	−9.26
F_1-O_3	−28.64	−29.82	−29.79	H_2-H_3	−6.16	−35.32	−5.93
F_1-H_4	−0.23	−0.22	−0.22	H_2-H_4	−5.16	1.88	−5.04
H_2-O_3	−32.04	−32.78	−32.67	H_2-N_5	−0.22	−0.22	−0.22
H_2-H_4	−0.23	−0.23	−0.23	H_4-N_5	−20.79	−21.40	−21.28
O_3-H_4	−528.95	-5.93×10^{5}	−477.77	H_4-H_6	−0.24	−0.24	−0.24
H_4-H_5	−2.11	−4.02	−1.96	H_4-H_7	−0.20	−0.20	−0.20
		HF–HF		N_5-H_6	−771.52	7.64×10^{5}	−702.05
F_1-H_2	−329.20	7048.78	−302.00	N_5-H_7	−770.75	6.85×10^{5}	−700.44
F_1-F_3	−17.02	−17.83	−17.82	H_6-H_7	−5.80	−22.66	−5.57
H_2-F_3	−16.71	−17.34	−17.32	H_7-H_8	−5.75	−21.54	−5.52
H_2-H_4	−0.11	−0.11	−0.11				
F_3-H_4	−356.83	-2.05×10^{5}	−320.36				
		HF–N_2					
F_1-H_2	−349.90	-5.85×10^{3}	−319.67				
F_1-N_3	−9.19	−9.58	−9.57				
F_1-N_4	−0.28	−0.28	−0.28				
H_2-N_3	−11.15	−11.46	−11.46				
H_2-N_4	−0.46	−0.45	−0.45				
N_3-N_4	−2266.80	-8.61×10^{5}	−2479.99				

valence orbitals) with the RHF results. All other *xc* interactions except those of the above CAS calculation lack a correlation energy component, being thus pure exchange contributions that should be more properly labelled V_x^{AB}. However, as all the expressions in section are valid for general wavefunctions the original name will be used hereinafter. The sums over l_1 and l_2 in eq. (15) were truncated at $l_1^{max}=l_2^{max}=8$, so that terms up to a range $L=l_1^{max}+l_2^{max}+1=17$ were included in the multipolar expansion. As QTAIM domains are usually finite and quite irregular, very fine radial and angular grids are needed to carry out the 6D numerical integrations. Here, we have systematically considered a $\beta-$sphere around each atom, with a radius equal to 60–90% the distance of its nucleus to the closest bond critical point, and used high quality Lebedev angular and radial grids, with (5810, 512) and (194, 400) points outside and inside the spheres, respectively. The errors in the total energy of the studied molecules attributable to these numerical integrations, necessarily approximate, are of the order of 1.0 kJ/mol. Our accumulated experience in IQA calculations makes us believe that the accuracy achieved in each interatomic interaction is even higher. Despite this issue regarding the full numerical accuracy of our integrations, once

Table 2. Representative *xc* interaction energies (kJ/mol) for the systems of Figure 3.

A − *B*	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	*A* − *B*	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
	Eclipsed BH_3-NH_3					**Li_9H_9**	
B_1-N_2	−164.57	3524.65	−172.41	Li_1-Li_2	−0.30	−0.31	−0.31
B_1-H_3	−355.03	8.40×10^6	−334.69	Li_1-Li_6	−0.71	−0.72	−0.72
B_1-H_6	−1.15	−1.14	−1.14	Li_1-H_{10}(1)	−17.22	−318.57	−17.75
N_2-H_3	−60.32	−62.95	−65.25	Li_1-H_{11}(1)	−25.92	42.71	−26.24
N_2-H_6	−723.06	-8.05×10^7	−659.17	Li_1-H_{15}	−0.47	−0.49	−0.49
H_3-H_4	−67.52	−339.26	−67.39	Li_2-Li_3	−0.70	−0.93	−0.71
H_3-H_6	−2.58	−4.71	−2.55	Li_2-Li_4	−0.14	−0.14	−0.14
H_3-H_7	−0.69	−0.68	−0.69	Li_2-Li_6	−0.84	−0.87	−0.85
H_6-H_7	−4.30	−865.54	−4.12	Li_2-H_{10}(1)	−26.95	331.98	−27.12
	Staggered BH_3-NH_3			Li_2-H_{11}(1)	−18.34	59.37	−18.74
B_1-N_2	−172.72	3080.97	−181.51	Li_2-H_{12}	−0.47	−0.50	−0.49
B_1-H_3	−351.06	7.76×10^6	−331.51	Li_2-H_{13}	−0.09	−0.10	−0.09
B_1-H_6	−1.21	−1.99	−1.20	Li_2-H_{15}(1)	−36.35	1.50×10^4	−36.27
N_2-H_3	−64.56	−71.22	−69.56	Li_2-H_{17}	−0.17	−0.17	−0.17
N_2-H_6	−721.25	-6.056×10^8	−659.38	Li_6-Li_7	−0.31	−0.31	−0.31
H_3-H_4	−66.25	−296.87	−66.17	Li_6-H_{10}	−0.84	−0.88	−0.88
H_3-H_6	−1.06	−1.17	−1.05	Li_6-H_{11}(1)	−36.95	36.59	−36.67
H_3-H_7	−1.62	−1.61	−1.61	Li_6-H_{12}	−0.19	−0.19	−0.19
H_6-H_7	−4.28	−472.57	−4.12	Li_6-H_{15}(1)	−45.11	-1.21×10^3	−44.54
		N_5^+		Li_6-H_{16}	−0.21	−0.21	−0.21
N_1-N_2	−1019.33	-1.70×10^5	−1056.45	H_{10}-H_{11}	−18.30	−20.78	−19.85
N_1-N_4	−47.28	−48.48	−48.20	H_{10}-H_{15}	−29.91	−38.97	−30.92
N_2-N_3	−52.15	−62.35	−55.59	H_{11}-H_{12}	−24.03	−36.36	−24.76
N_2-N_4	−2066.45	2.90×10^7	−2205.41	H_{11}-H_{13}	−0.70	−0.71	−0.70
N_2-N_5	−7.78	−7.77	−7.77	H_{11}-H_{15}	−25.73	−49.73	−26.56
N_4-N_5	−3.26	−3.27	−3.27	H_{11}-H_{17}	−0.28	−0.28	−0.28
				H_{15}-H_{16}	−1.48	−1.47	−1.44
				H_{15}-H_{17}	−0.19	−0.19	−0.19

In Li_9H_9, the directly bonded Li and H atoms are signalled with (1).

the computational conditions of a given calculation have been chosen the convergence of the bipolar expansion (the exact benchmark) is ordinarily well below the 1 kJ/mol barrier for the *xc* contributions.

Results and Discussion

The more representative results regarding the approximate V_{xc}^{AB} values, as well as their errors for the systems listed in Section are gathered in Tables 1–3 and Figures 7–11. We can see in Table 1, where the V_{xc}^{AB}'s for the HB systems of Figure 2 are collected, that the full MP $(V_{xc}^{AB})_{lr}$ [eq. (15)] fails miserably for all intramolecular A–H pairs (A=N,O,F). Surprisingly, the crude (lr,cdq) approximation gives *xc* interactions with relative errors of about 10% or smaller for the intramolecular directly bonded atoms. Regarding the intermolecular interactions, the *xc* energy between the two atoms involved in the HB is well represented by $(V_{xc}^{AB})_{lr}$, with differences with respect to the exact values smaller than 0.3 kJ/mol in all the cases. We note again that the (lr,cdq) values differ only by 0.3–0.6 kJ/mol from the exact ones, confirming that the multipolar expansion for these interactions is practically converged at this level of calculation. Intermolecular A–H and H–H *xc* energies other than the above ones are given by the lr approximation with errors smaller than 0.1 and 0.01 kJ/mol, respectively. For the intermolecular H–H energies, the same is true in the (lr,cdq) approximation. However, the *xc* interaction between the A atom of the proton donor (PD) and the *B* atom of the proton acceptor (PA) molecule is predicted with errors as large as 1.4 kJ/mol ($FH\cdots NH_3$) when the (lr,cdq) approximation is used, which clearly indicates that multipolar interactions higher than the charge-quadrupole ones included in this approximation are required to represent this type of interaction with accuracy.

The relative errors of the A–B, A–H (A,B=N,O,F), and H–H *xc* interaction energies for all the intramolecular and intermolecular pairs of the HB systems are represented in Figure 7. We observe in Figure 7 that cd and cdq intramolecular H–H energies are, in general, more accurate than the ones for the A–H interactions, which is clearly due to the 1–3 (1–2) character of all the intramolecular H–H (A–H) pairs. It is also striking that, with a couple of exceptions, cdq relative errors are negative whereas the contrary happens with the cd approximation. Moreover, as previously commented, only a single lr relative error appears in the figure, the remaining ones having errors greater than 20%. Regarding the intermolecular *xc* energies, we observe in right Figure 7 the progressive decreasing of relative errors in passing from cd to cdq, and from cdq to lr.

The discussion for the systems in Figure 3 runs parallel to that of the HB dimers. The $(V_{xc}^{AB})_{lr}$ value for the B–N pair in eclipsed and staggered BH_3-NH_3 has no sense. Similarly, the lr *xc* interaction between the directly bonded (i.e., 1–2) B–H and N–H pairs is quite absurd. Not only that, but also the $(V_{xc}^{AB})_{lr}$'s for the 1–3 pairs H_3-H_4 and H_6-H_7 are several orders of magnitude greater than the exact values. Contrarily to this,

Table 3. Representative *xc* interaction energies (kJ/mol) for the systems of Figure 4.

A − *B*	$(V_{xc}^{AB})_{Ir,cdq}$	$(V_{xc}^{AB})_{Ir}$	V_{xc}^{AB}	*A* − *B*	$(V_{xc}^{AB})_{Ir,cdq}$	$(V_{xc}^{AB})_{Ir}$	V_{xc}^{AB}
		***cis*-$C_2H_2F_2$**				**CH_3BeH**	
C_1-C_2	−1311.48	-3.82×10^5	−1351.45	H_1-C_2	−34.84	−32.52	−32.62
C_1-H_3	−811.14	-1.20×10^4	−746.88	H_1-H_3	−0.75	−0.74	−0.75
C_1-H_4	−17.00	−15.11	−15.79	H_1-Be_6	−170.49	2.43×10^4	−159.81
C_1-F_5	−600.29	-5.29×10^3	−608.87	C_2-H_3	−845.84	6.99×10^5	−765.34
C_1-F_6	−53.21	−54.86	−52.56	C_2-Be_6	−168.87	−824.46	−164.53
H_3-H_4	−1.40	−1.40	−1.40	H_3-H_4	−17.84	−30.31	−17.43
H_3-F_5	−24.53	−24.79	−24.54	H_3-Be_6	−2.46	−2.45	−2.47
H_3-F_6	−3.64	−3.70	−3.52			**CH_3BH_2**	
F_5-F_6	−4.61	−4.60	−4.60	C_1-B_2	−361.03	−496.34	−362.60
		trans-$C_2H_2F_2$		C_1-H_3	−826.73	2.48×10^5	−745.50
C_1-C_2	−1304.91	-3.87×10^5	−1345.33	C_1-H_4	−834.27	6.39×10^4	−762.61
C_1-H_3	−811.92	−540.92	−747.51	C_1-H_6	−66.37	−68.63	−68.96
C_1-H_4	−16.93	−15.04	−15.72	B_2-H_3	−6.72	−6.70	−6.71
C_1-F_5	−603.33	-5.67×10^3	−611.80	B_2-H_4	−4.08	−4.11	−4.11
C_1-F_6	−54.31	−56.04	−53.69	B_2-H_6	−376.42	-1.85×10^4	−349.54
H_3-H_4	−1.37	−1.38	−1.38	H_3-H_4	−17.11	−25.73	−16.68
H_3-F_5	−24.82	−25.10	−24.84	H_3-H_6	−2.04	−2.04	−2.04
H_3-F_6	−3.66	−3.72	−3.53	H_4-H_5	−15.65	−21.75	−15.36
F_5-F_6	−4.71	−4.69	−4.69	H_4-H_6	−2.82	−2.81	−2.81
		CH_2FOH		H_4-H_7	−2.11	−2.10	−1.72
C_1-O_2	−595.20	1.19×10^3	−613.00	H_6-H_7	−66.83	−69.51	−65.42
C_1-F_3	−515.50	80.74	−529.59			**CH_3CH_2F**	
C_1-H_4	−784.72	-1.93×10^4	−724.78	C_1-C_2	−776.25	−808.13	−790.31
C_1-H_6	−3.15	−3.08	−3.12	C_1-F_3	−34.45	−36.53	−36.63
O_2-F_3	−84.54	−94.13	−94.16	C_1-H_4	−818.55	5.40×10^4	−748.07
O_2-H_4	−29.48	−29.36	−29.61	C_1-H_5	−820.78	7.25×10^4	−749.52
O_2-H_6	−511.47	-2.86×10^5	−465.17	C_1-H_7	−14.33	−14.45	−14.31
F_3-H_4	−31.06	−31.02	−31.26	C_2-F_3	−546.94	10.18	−558.24
F_3-H_6	−5.11	−5.18	−5.19	C_2-H_4	−17.73	−17.99	−17.87
H_4-H_5	−11.89	−13.77	−11.64	C_2-H_5	−15.64	−15.90	−15.82
H_4-H_6	−0.47	−0.47	−0.47	C_2-H_7	−804.65	2.30×10^5	−740.40
		CH_3CF_3		F_3-H_4	−2.92	−2.90	−2.90
C_1-H_2	−818.30	8.69×10^4	−746.71	F_3-H_5	−3.47	−3.54	−3.54
C_1-C_5	−735.49	−846.24	−758.72	F_3-H_7	−31.02	−31.13	−31.29
C_1-F_6	−31.84	−33.40	−33.56	H_4-H_5	−15.48	−21.10	−15.16
H_2-H_3	−13.27	−16.44	−13.01	H_4-H_7	−1.07	−1.06	−1.06
H_2-C_5	−16.37	−16.79	−16.80	H_5-H_6	−15.14	−20.32	−14.82
H_2-F_6	−2.55	−2.53	−2.53	H_5-H_7	−1.18	−1.17	−1.17
H_2-F_7	−3.45	−3.53	−3.21	H_5-H_8	−2.08	−2.08	−4.90
C_5-F_6	−469.52	−792.76	−488.98	H_7-H_8	−13.54	−17.74	−13.22
F_6-F_7	−83.69	−93.89	−93.53				
C_1-C_2	−789.95	9.49×10^3	−798.97	C_1-N_2	−740.43	1.21×10^3	−753.96
C_1-Li_3	−1.66	−1.67	−1.67	C_1-H_3	−806.63	1.18×10^4	−740.54
C_1-H_4	−810.07	1.51×10^5	−743.11	C_1-H_4	−809.44	9.59×10^4	−741.04
C_1-H_5	−815.31	2.79×10^5	−746.34	C_1-H_6	−6.68	−7.10	−6.60
C_1-H_7	−21.37	−22.48	−21.68	N_2-H_3	−29.22	−29.36	−28.96
C_2-Li_3	−100.43	−554.91	−96.79	N_2-H_4	−27.42	−27.66	−27.76
C_2-H_4	−19.31	−36.64	−18.43	N_2-H_6	−770.22	1.45×10^6	−707.83
C_2-H_5	−18.60	−13.76	−18.54	H_3-H_4	−16.11	−21.94	−15.75
C_2-H_7	−838.53	3.78×10^6	−767.82	H_3-H_6	−0.59	−0.58	−0.58
Li_3-H_4	−0.57	−0.57	−0.57	H_4-H_5	−15.88	−22.49	−15.48
Li_3-H_5	−0.16	−0.16	−0.16	H_4-H_6	−0.62	−0.61	−0.61
Li_3-H_7	−1.98	−1.98	−1.96	H_4-H_7	−1.91	−1.90	−1.77
H_4-H_5	−18.94	−36.17	−18.51	H_6-H_7	−5.87	−16.69	−5.67
H_4-H_7	−1.59	−1.57	−1.57			**CH_3OH**	
H_5-H_6	−19.74	−36.04	−19.25	C_1-O_2	−638.41	−187.46	−650.07
H_5-H_7	−1.26	−1.25	−1.25	C_1-F_3	−808.68	-6.14×10^3	−740.99
H_5-H_8	−2.65	−2.64	−2.64	C_1-H_4	−804.40	2.53×10^4	−738.45
H_7-H_8	−23.12	−97.48	−22.57	C_1-H_6	−3.79	−3.81	−3.78
		CH_3CLi_3		O_2-F_3	−36.31	−36.22	−36.71
C_1-H_2	−806.90	4.89×10^5	−739.52	O_2-H_4	−29.92	−29.73	−30.14
C_1-C_5	−823.71	2.89×10^4	−839.98	O_2-H_6	−553.82	-3.09×10^5	−503.97
C_1-Li_6	−2.38	−2.42	−2.41	F_3-H_4	−15.27	−21.13	−14.92
H_2-H_3	−20.16	−63.55	−19.70	F_3-H_6	−1.32	−1.31	−1.32
H_2-C_5	−25.16	−33.67	−25.04	H_4-H_5	−14.69	−18.89	−14.35

(Continued)

Table 3. (*Continued*)

$A-B$	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	$A-B$	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
H_2-Li_6	−0.65	−0.64	−0.64	H_4-H_6	−0.50	−0.50	−0.50
H_2-Li_7	−0.43	−0.43	−0.38			C_6H_6	
C_5-Li_6	−135.99	-1.03×10^5	−123.94	C_1-C_2	−1043.99	-2.45×10^6	−1065.06
Li_6-Li_7	−1.34	−3.02	−1.29	C_1-C_3	−23.85	−26.79	−24.50
				C_1-C_4	−22.49	−23.40	−23.40
				C_1-H_7	−819.10	1.90×10^7	−756.10
				C_1-H_8	−16.89	−18.55	−17.07
				C_1-H_9	−1.73	−1.73	−1.73
				C_1-H_{10}	−0.92	−0.93	−0.93
				H_7-H_8	−1.84	−1.86	−1.84
				H_7-H_9	−0.21	−0.21	−0.21
				H_7-H_{10}	−0.05	−0.05	−0.05

the cdq approximation works relatively well for the B–N pair and the 1–2 B–H and N–H pairs (relative error <5%). The H_3-H_4 interaction is also extremely well reproduced by this approximation (error <0.2%), whereas the H_6-H_7 is slightly worse (error ~4%). The *xc* interaction between the B atom and a H atom of the NH_3 unit (B_1-H_6) is fairly accurate in both the lr and cdq approximations. This is not so with the symmetric interaction N_2-H_3, with errors about 4–7% in both cases.

In the N_5^+ molecule, the *xc* energy for the 1–2 pairs is again badly represented by the lr approximation but is reasonable in the cdq approach, particularly for N_1-N_2. The cdq and lr values for the 1–3 N_1-N_4 interaction differs from the exact value by about 0.9 and 0.3 kJ/mol, respectively. The error in the other 1–3 interaction (N_2-N_3) in considerably higher in both approaches. Finally, the lr and cdq *xc* energies for the 1–4 N_2-N_5 and 1–5 N_4-N_5 pairs are practically the same and coincident with the exact value. This result highlights two important facts: (i) the atomic basins of N_2 and N_5 (or N_4 and N_5) atoms fullfil almost exactly the non-overlapping criterion displayed in Figure 1, and (ii) the multipolar series 15 converges very quickly in this particular case.

Finally, the results for Li_9-H_9 reinforce what was said in the above three paragraphs. The full lr expansion fails completely in predicting *xc* interaction energies for 1–2 Li–H pairs, while the cdq values are pretty accurate. All Li–Li *xc* energies are well represented in the lr and cdq approximations, with the exception of the lr Li_2-Li_3 interaction. This is probably related with the almost spherical character of Li atomic basins. According to this, the 1–3 Li–H lr *xc* energies and, more importantly, the 1–3 H–H *xc* energies are less accurately computed due to the far from the spherical character of H atomic basins. This is exacerbated in the lr approximation, where higher angular number *l* values are involved [see eq. (17)].

The *xc* pair interaction energies of the systems in Figure 4 are collected in Table 3, and the relative errors of the cd, cdq and lr approximate values displayed in Figure 8, for the 1–2 (left-top), 1–3 (right-top), and 1–4 (bottom) pairs, respectively. Virtually all of the above comments also apply here: the 1–2 *xc* interactions cannot be represented at all using the full lr

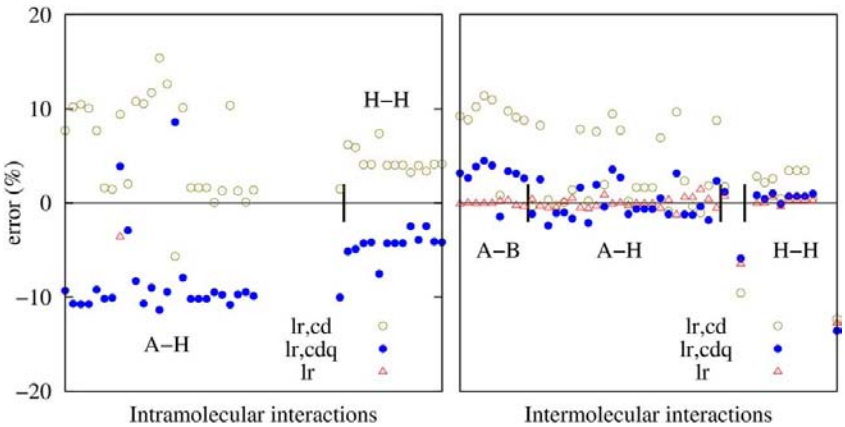


Figure 7. Relative errors, $[(V_{xc}^{AB})_{method}-(V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}|\times100$, of the intramolecular (left) and intermolecular (right) interactions of the molecules in Figure 2. [Color figure can be viewed at wileyonlinelibrary.com]

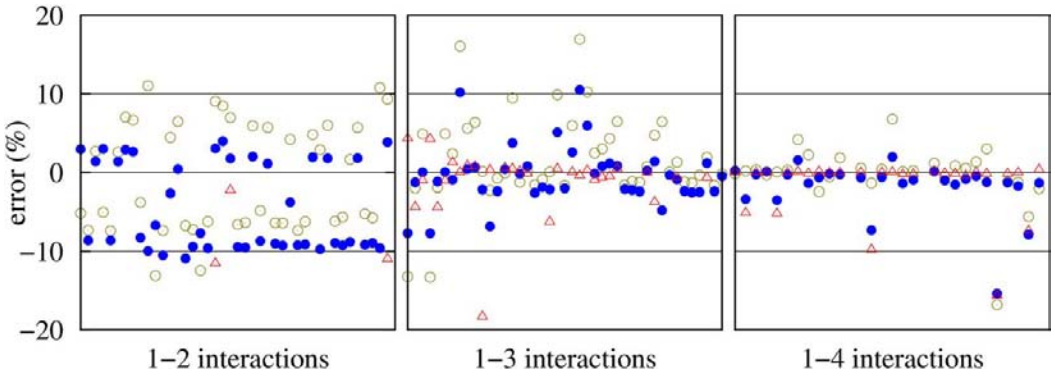


Figure 8. Relative errors, $[(V_{xc}^{AB})_{method}-(V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}|\times 100$, for the 1−2, 1−3, and 1−4 interactions of the molecules represented in Figure 4. Empty circles, bold circles, and triangles stand for (lr,cd), (lr,cdq), and lr calculations, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

expansion. However, they are given with reasonable accuracy by the cdq approximation. The 1 – *n* ($n>2$) interactions are gradually better reproduced as *n* increases in both the lr and cdq approximations. It is very satisfactory to check that the cdq approach, a severe truncation of the full *mp* expansion, is perfectly suited to simulate the *xc* interaction between pairs of atoms beyond the directly bonded ones. Even in typically covalent molecules like benzene all the cdq C–C *xc* interaction energies reproduce very well the exact values. As we can see in Figure 8 most of the 1–2 interactions have relative cdq errors $\leq$ 10%. This improves for the 1–3 and 1–4 interactions.

A summary of our results for the saturated hydrocarbons C_nH_{2n+2} ($n=2-5$) is presented in graphical form in Figure 9. We find the surprising result that all 1–2 C–C cdq interactions are predicted with errors $\leq$ 2% while the cdq energies between the more distant 1–3 C–C pairs have errors about 5–6%. Nevertheless, the interactions between even more distant C–C pairs turn again to be calculated quite accurately (errors <1%) in the cdq approximation. The lr approximation fails completely to predict the 1–2 C–C interactions, but yields negligible errors for the 1–3 *xc* interaction energies. With regard to the C–H interactions, the situation is the opposite of that found for the C–C pairs: 1–2 C–H cdq errors are about 9–10% (except in ethane where the error is unusually large (54%)) whereas all except two of the 1–3 C–H cdq errors are <1%. For these two exceptions the error is not too large (∼1.3%). We observe in Figure 9 that the cdq approximation improves considerably the cd results, giving 1–3 C–H interaction energies almost as accurate as the lr ones. Another surprising result in these systems concerns the cdq 1–3 H–H interactions: Contrary to what happens almost systematically, the cdq results are worse than the cd ones, albeit the relative errors in both approximations are acceptable (∼2−3%).

We have considered staggered ethane as a representative example to analyze the correlation effects on the V_{xc}^{AB} energies. Our results for five representative AB pairs of this molecule are collected in Table 4. Correlation decreases (increases) the magnitude of the C–C interaction (H–H interactions), changes very

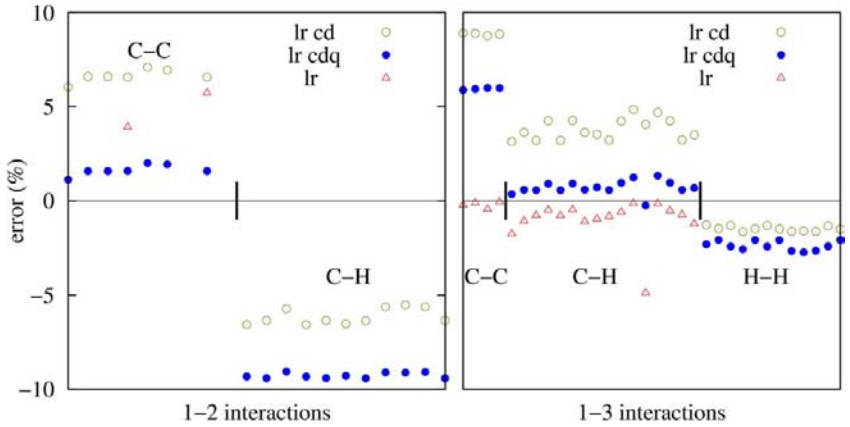


Figure 9. Relative errors, $[(V_{xc}^{AB})_{method}-(V_{xc,exact}^{AB})]/|V_{xc,exact}^{AB}|\times 100$, for the 1−2 (left) and 1−3 (right) interactions of the molecules represented in Figure 5. [Color figure can be viewed at wileyonlinelibrary.com]

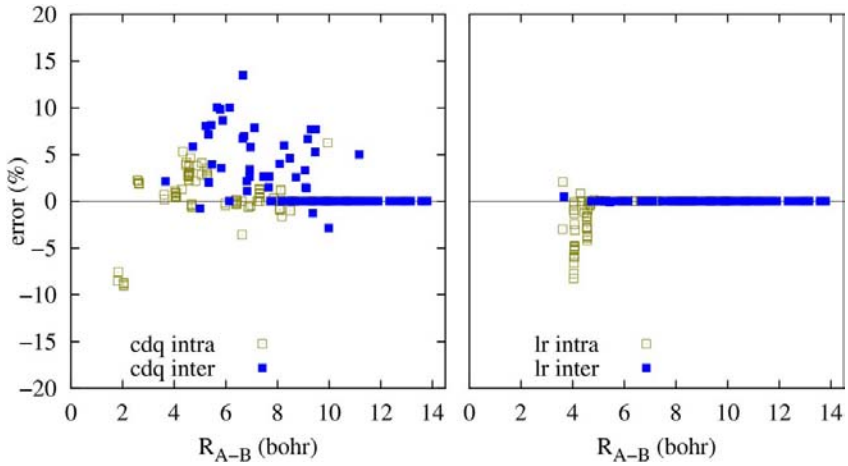


Figure 10. Relative error of the cdq (left) and lr (right) calculations for the phenol-dimer ($C_6H_5OH\cdots C_6H_5OH$). Only two (2) points are out of the ordinate scale in the cdq calculation, while all the points (26) with $R_{A-B} < 2.64$ bohr are out of the ordinate scale in the lr calculation. [Color figure can be viewed at wileyonlinelibrary.com]

little V_{xc}^{CH} when C belongs to a CH_3 group and H to the other, and enhances the intra-group V_{xc}^{CH} energies. The discussion of the above paragraph for the RHF results is still approximately valid for the correlated calculation. With the exception of inter-group H–H interactions, the cdq *xc* energies are closer to the exact V_{xc}^{AB} values than their full MP $(V_{xc}^{AB})_{lr}$. As in many of the HF 1–2 interactions, the lr approximation fails to predict even a reasonable value for the C_1-H_5 *xc* energy.

The different behavior of the lr and cdq approximations can be further illustrated with the case of the phenol dimer (Fig. 6). For this system, the relative errors *versus* the interatomic distance R_{A-B} in these two approximations are plotted in Figure 10, both for intramolecular and intermolecular atomic pairs. Only two points, associated to intramolecular interactions, have a relative error (absolute value) $\geq 20\%$ in the cdq calculation, while the error for all the lr points with $R_{A-B} < 2.64$ (most of them associated to intramolecular pairs) is larger than 20%. However, for $R_{A-B} > 5.0$ the lr approximation gives quite accurate *xc* interaction energies for all the pairs, whereas cdq errors are still important.

However, there is a general problem of the lr approximation that deserves to be commented: eq. (15) does not necessary converges to the exact *xc* interaction for large l_1m_1 and l_1m_1 values. This fact is illustrated in Figure 11, where the *xc* energies for some of the atomic pairs of the molecules in Figure 4 are represented versus $l_1^{max}+l_2^{max}$. The lr approximation suffers a systematic error in the C_1-F_6 interaction of cis-CH_2CF_2 and trans-CH_2CF_2 molecules, regardless the value of $l_1^{max}+l_2^{max}$. In the case of trans-CH_2CF_2, the C_1-F_6 *xc* interaction energy shows an oscillating behavior around an (erroneous) mean

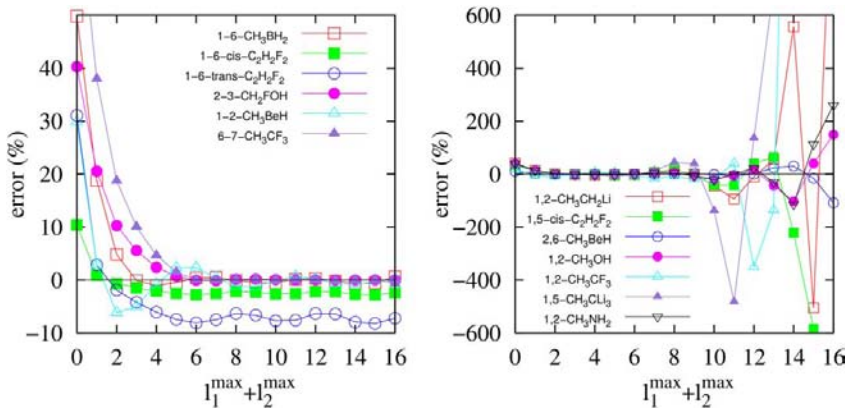


Figure 11. Convergence of the *A* – *B* interactions indicated in the figure. The pairs *i*, *j* correspond to the labels of Figure 4. [Color figure can be viewed at wileyonlinelibrary.com]

Table 4. Comparison of RHF and CAS[14,14] *xc* energies (kJ/mol) for staggered C_2H_6.

	RHF			CAS[14,14]		
	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}	$(V_{xc}^{AB})_{lr,cdq}$	$(V_{xc}^{AB})_{lr}$	V_{xc}^{AB}
C_1-C_2	−781.87	−870.42	−790.68	−625.88	−723.66	−639.75
C_1-H_5	−821.98	8.93×10^4	−533.05	−686.06	8.17×10^4	−618.64
H_3-H_4	−16.75	−24.11	−16.38	−25.06	−32.82	−24.75
C_1-H_3	−16.21	−16.55	−16.26	−16.62	−16.75	−16.55
H_3-H_6	−2.32	−2.32	−2.32	−3.10	−3.10	−3.10

value. This pattern has been also observed in other cases. Contrarily, as we have repeatedly said in this section, catastrophic lr interactions (see, for instance Fig. 11) are still reasonable provided that the sum l_1+l_2 is interrupted at a value approximately in the interval $2 \leq l_1+l_2 \leq 6$.

This simple analysis shows that conclusions on the convergence of the *mp* expansion drawn exclusively from limited $L=l_1^{max}+l_2^{max}+1$ data cannot be trusted. Figure 12, where the $(V_{xc}^{AB})_{lr}$ energies for some 1−2 interactions of the H_2O–H_2O, HF-HF, and NH_3-NH_3 dimers are shown, illustrates this fact in a crystal clear way. Cutting the *mp* expansion at $L \leq 8$ the oscillatory behavior of Figure 11 would have been found indeed, but no catastrophe for larger *L* values would have been predicted. However, for $L>8$ the *mp* expansion progressively deteriorates and for high *L* values the $(V_{xc}^{AB})_{lr}$ energies diverge. Again, the cdq approximation works fairly well, with all the 1–2 interactions in Figure 12 predicted with relative errors smaller than 3% except F_1-H_2 (3.2%) and F_3-H_4 (5.2%).

Summarizing, we have shown that the multipole series for the interatomic *xc* energies is conditionally convergent, and that the computational burden of the quasi-exact calculation of V_{xc} when general QTAIM domains are used may be ameliorated by retaining up to quadrupole-quadrupole terms. With this approximation, reasonable errors are obtained in the medium- to long-distance range, sometimes even for directly bonded interactions. An overall image of the improvement of the cdq approximation over the cc (monopole-monopole) one can be grasped from Figure 13 that condenses all our calculations that span a six orders of magnitude range for V_{xc}.

We have explored numerically the degree of fulfillment of the linear relation between V_{xc}^{AB} and $\delta^{AB}/(2R)$. Assuming $V_{xc}^{AB} \simeq -a[\delta^{AB}/(2R)]+b$, we have determined *a* and *b* for each molecule by fitting the computed values of V_{xc}^{AB} and δ^{AB} for every AB pair of this molecule to the above expression, verifying that *b* is always very small and *a* takes values relatively close to (but smaller than) 1.0. As representative examples, (*a*, *b*) for the H_2O-H_2O, NH_3-NH_3, and HF-HF dimers are $(a,b)=(0.8988,-0.0005)$, $(0.8673,0.0000)$, and $(0.9150,-0.0010)$, respectively. This result opens another possible route to the approximate but much cheaper computation of the *xc* interaction in those cases where the exact calculation is prohibitive or very expensive. An extensive analysis of this correlation in anion-π interactions which corroborates the above statement has been carried out by Foroutan-Nejad et al.[37–39]

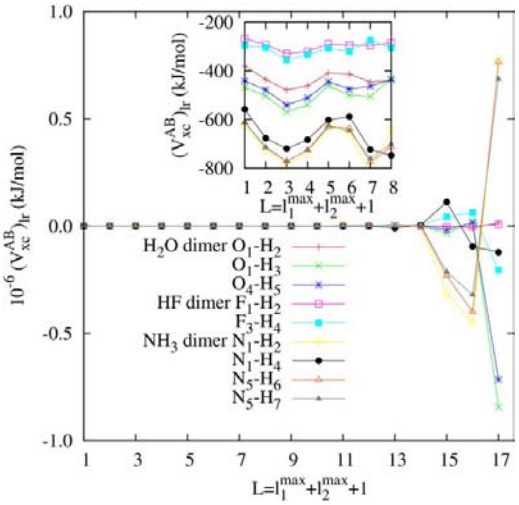


Figure 12. Convergence of 1−2 interactions in the H_2O–H_2O, HF-HF, and NH_3-NH_3 dimers. [Color figure can be viewed at wileyonlinelibrary.com]

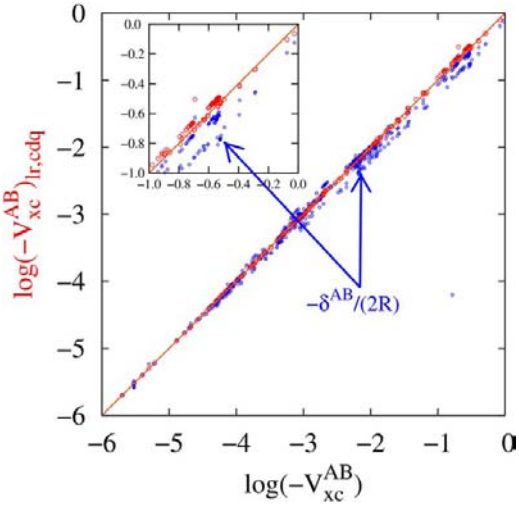


Figure 13. Comparison of all the exact V_{xc}^{AB} values considered in this work to the monopole-monopole (small circles) and cdq approximations (large circles) in a logarithmic scale. [Color figure can be viewed at wileyonlinelibrary.com]

Conclusions

We have shown that the interatomic exchange-correlation energies used in real space theories of chemical bonding, which measure the covalent contribution to a given interatomic interaction, can be approximated via a conventional multipole expansion. Rigorously, the series diverges when atoms are directly bonded, although it may be regarded asymptotically convergent. Truncation of the series up to $l_1+l_2=2$ (including up to charge–quadrupole interactions) tends to provide results which are accurate to a few percent in $1-n$, $n>2$ interactions, and even to about 10% in many 1–2 directly bonded cases. In the $n>2$ case the series converges in many cases, and including extra terms provides further accuracy. On the contrary, the consideration of larger l contributions in 1–2 interactions tends to seriously deteriorate the results. As the computational burden needed to calculate the multipole series is considerably smaller than that of the exact bipolar expansion, our results may be important to estimate covalent interactions in those cases where exact integrations are not feasible. They can also be used to ameliorate the computational cost in IQA decompositions of large systems, where many expensive, but small long-range *xc* terms can now be safely approximated without loss of precision.

Appendix

We derive in this appendix eq. (15), the MP to the exact exchange–correlation interaction, eqs. (10) and (11). Further details are given in I. We start using the bipolar expansion for r_{12}^{-1},

$$r_{12}^{-1}=\sum_{l_1m_1}^{\infty}\sum_{l_2m_2}^{\infty} S_{l_1m_1}(\hat{r}_1)\, S_{l_2m_2}(\hat{r}_2)\, D_{l_1m_1}^{l_2m_2}(r_1,r_2,\mathbf{R}), \tag{24}$$

where $\mathbf{r}_1\equiv(r_1,\hat{r}_1)$ and $\mathbf{r}_2\equiv(r_2,\hat{r}_2)$ are referred to centers A and B, respectively, $\mathbf{R}=(\mathbf{R}_B-\mathbf{R}_A)\equiv(R,\hat{R})$ is the position vector

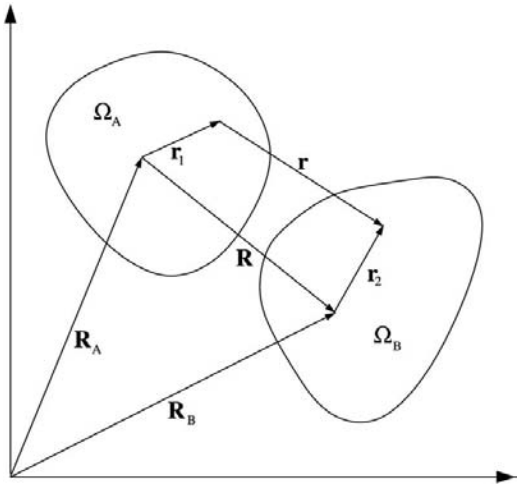


Figure 14. Coordinate system.

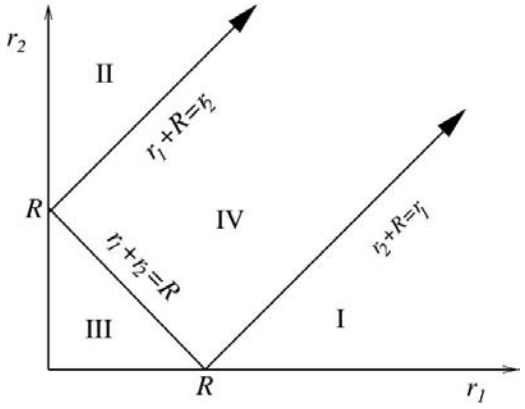


Figure 15. Regions of definition of the $\mathcal{V}_{l_1l_2l_3}(r_1,r_2,R)$ function.

of center B with respect to center A (see Fig. 14), $S_{lm}(\hat{r})$ are real spherical harmonics defined as[32]

$$S_{lm}(\theta,\phi)=\Theta_{l|m|}(\theta)\Phi_m(\phi), \tag{25}$$

$$\Theta_{lm}(\theta)=\sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}P_l^m(\cos\theta), \tag{26}$$

$$\Phi_m(\phi)=\begin{cases}\sqrt{2}\cos m\phi & m>0,\\ 1 & m=0,\\ \sqrt{2}\sin|m|\phi & m<0,\end{cases} \tag{27}$$

and $P_l^m(\cos\theta)$ are the associated Legendre functions, defined for $m\geq 0$ by

$$P_l^m(x)=\frac{1}{2^l l!}(1-x^2)^{m/2}\frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l. \tag{28}$$

Finally, $D_{l_1m_1}^{l_2m_2}(r_1,r_2,\mathbf{R})$ in eq. (24) is defined as

$$D_{l_1m_1}^{l_2m_2}(r_1,r_2,\mathbf{R})=4\pi(-1)^{l_1}\sum_{l_3=|l_1-l_2|}^{l_1+l_2}\mathcal{V}_{l_1l_2l_3}(r_1,r_2,R)\,T_{l_1m_1l_2m_2}^{l_3}(\hat{R}), \tag{29}$$

where the sum over l_3 runs in steps of 2, $\mathcal{V}_{l_1l_2l_3}(r_1,r_2,R)$ is a discriminant that takes different expressions in the four regions defined in Figure 15, and $T_{l_1m_1l_2m_2}^{l_3}(\hat{R})$ is the angular factor

$$T_{l_1m_1l_2m_2}^{l_3}(\hat{R})=\sum_{m_3=-l_3}^{+l_3} d_{l_1m_1l_2m_2}^{l_3m_3}\,S_{l_3m_3}(\hat{R}), \tag{30}$$

where $d_{l_1m_1l_2m_2}^{l_3m_3}$ is the Gaunt coefficient between the $S_{lm}(\theta,\phi)$'s defined by

$$d_{l_1m_1l_2m_2}^{l_3m_3}=\langle S_{l_3m_3}|S_{l_1m_1}|S_{l_2m_2}\rangle. \tag{31}$$

Given that S_{lm} is real, $d_{l_1m_1l_2m_2}^{l_3m_3}$ is invariant against any permutation of the pair of indices (l_i, m_i). These coefficients may be determined as described elsewhere. Using eq. (24) in eq. (11) one has

$$K_{ij}^{AB}=\sum_{l_1m_1}\sum_{l_2m_2}\int_{\Omega_A} S_{l_1m_1}(\hat{r}_1)\, f_{ij}(\mathbf{r}_1)\, d\mathbf{r}_1 \int_{\Omega_B} S_{l_2m_2}(\hat{r}_2)\, f_{ij}(\mathbf{r}_2)\, d\mathbf{r}_2\, D_{l_1m_1}^{l_2m_2}(r_1,r_2,\mathbf{R})). \quad (32)$$

A further simplification of K_{ij}^{AB} requires the explicit form of $D_{l_1m_1}^{l_2m_2}(r_1,r_2,\mathbf{R})$. From the expression of $\mathcal{V}_{l_1l_2l_3}(r_1,r_2,R)$ [eqs. (B1)–(B9) of I] it follows that, as long as $r_1+r_2 \le R$, this discriminant only takes a nonzero value in region III of the (r_1, r_2) space (see Fig. 15). This condition will be exactly satisfied if $R \ge r_1^{max}+r_2^{max}$, where r_1^{max} is the maximum value of the radial coordinate within Ω_A, with an equivalent definition for r_2^{max}. In the present context, atoms A and B are said to be non-overlapping if this condition is fulfilled, and overlapping otherwise. Although it may occur that the condition $R \ge r_1^{max}+r_2^{max}$ is not exactly satisfied, provided that the atomic basins Ω_A and Ω_B are well-separated in the space, we can expect that it is fulfilled in practical terms. The multipolar approach, intensively used to approximate the Coulomb repulsion in the modelization of biomolecules, is equivalent to the assumption that $r_1+r_2 \le R$ for any r_1 and r_2. Thus, region III is identified with the complete first quadrant. In this region, $D_{l_1m_1}^{l_2m_2}(r_1,r_2,\mathbf{R})$ is given by

$$D_{l_1m_1}^{l_2m_2}(r_1,r_2,\mathbf{R})=(-1)^{l_1}16\pi^2\,\Delta_{l_1l_2}\,\frac{r_1^{l_1}r_2^{l_2}}{R^{l_1+l_2+1}}\,T_{l_1m_1,l_2m_2}^{l_1+l_2}(\hat{R}), \text{where} \quad (33)$$

$$\Delta_{l_1l_2}=(-1)^{l_1+l_2}\frac{(2l_1+2l_2)!\,l_1!\,l_2!}{(l_1+l_2)!\,(2l_1+1)!\,(2l_2+1)!}. \quad (34)$$

Using eq. (33) in eq. (32), we get

$$\left(K_{ij}^{AB}\right)_{lr}=\sum_{l_1m_1}\sum_{l_2m_2} C_{l_1m_1l_2m_2}(\hat{R})\frac{q_{ij,l_1m_1}^{\Omega_A}\,q_{ij,l_2m_2}^{\Omega_B}}{R^{l_1+l_2+1}}, \text{where} \quad (35)$$

$$C_{l_1m_1,l_2m_2}(\hat{R})=(-1)^{l_1}4\pi[(2l_1+1)(2l_2+1)]^{\frac{1}{2}}\Delta_{l_1l_2}\,T_{l_1m_1l_2m_2}^{l_1+l_2}(\hat{R}), \quad (36)$$

and the $q_{ij,lm}^{\Omega}$ have been defined in eq. (17). Finally, substituting eq. (35) in eq. (10) we obtain eq. (15), the MP for the exchange–correlation interaction, $\left(V_{xc}^{AB}\right)_{lr}$.

Keywords: chemical bonding · atoms in molecules · covalent energy · molecular energy partitioning · electron delocalization

How to cite this article: E. Francisco, D. Menéndez Crespo, A. Costales, A. Martín Pendás. *J. Comput. Chem.* **2017**, DOI: 10.1002/jcc.24758

[1] S. Kurth, J. P. Perdew, *Int. J. Quantum Chem.* **2000**, *77*, 814.
[2] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133.
[3] E. J. Baerends, O. V. Gritsenko, *J. Phys. Chem. A* **1997**, *101*, 5383.
[4] E. Francisco, A. Martín Pendás, M. A. Blanco, *J. Chem. Theory Comput.* **2006**, *2*, 90.
[5] E. Francisco, A. Martín Pendás, M. A. Blanco, *J. Chem. Theory Comput.* **2006**, *2*, 90.
[6] A. Martín Pendás, M. A. Blanco, E. Francisco, *J. Comput. Chem.* **2007**, *28*, 161.
[7] M. Kohout, *Int. J. Quantum Chem.* **2004**, *97*, 651.
[8] M. Causà, A. Savin, *J. Phys. Chem. A* **2011**, *115*, 13139.
[9] B. M. Gimarc, *Molecular Structure and Bonding. The Qualitative Molecular Orbital Approach;* Academic Press: New York, **1979**.
[10] R. F. W. Bader, *Atoms in Molecules;* Oxford University Press: Oxford, **1990**.
[11] A. D. Becke, K. E. Edgecombe, *J. Chem. Phys.* **1990**, *92*, 5397.
[12] B. Silvi, A. Savin, *Nature* **1994**, *371*, 683.
[13] P. L. A. Popelier, E. A. Brèmond, *Int. J. Quantum Chem.* **2009**, *109*, 2542.
[14] A. Martín Pendás, E. Francisco, M. A. Blanco, *J. Phys. Chem. A* **2006**, *110*, 12864.
[15] J. M. Guevara-Vela, C. C. R. M. García-Revilla, J. Hernández Trujillo, O. Christiansen, E. Francisco, A. Martín Pendás, T. Rocha-Rinza, *Chem. Eur. J.* **2013**, *19*, 4304.
[16] C. Gatti, P. Macchi, Eds. *Modern Charge-density Analysis;* Springer: Dordrecht, **2012**.
[17] R. Chauvin, C. Lepetit, B. Silvi, E. Alikhani, Eds. *Applications of Topological Methods in Molecular Chemistry;* Springer Intl., Switzerland, **2016**.
[18] A. Martín Pendás, E. Francisco, M. A. Blanco, C. Gatti, *Chem. Eur. J.* **2007**, *13*, 9362.
[19] M. García-Revilla, P. L. A. Popelier, E. Francisco, A. Martín Pendás, *J. Chem. Theory Comput.* **2011**, *7*, 1704.
[20] E. Bartashevich, E. Troitskaya, A. Martín Pendás, V. G. Tsirelson, *Comput. Theor. Chem.* **2015**, *1053*, 229.
[21] E. Bartashevich, A. Martín Pendás, V. G. Tsirelson, *Phys. Chem. Chem. Phys.* **2014**, *16*, 16780.
[22] A. Martín Pendás, M. A. Blanco, E. Francisco, *J. Comput. Chem.* **2009**, *30*, 98.
[23] M. García-Revilla, E. Francisco, P. L. A. Popelier, A. Martín Pendás, *ChemPhysChem* **2013**, *14*, 1211.
[24] R. F. W. Bader, M. E. Stephens, *Chem. Phys. Lett.* **1974**, *26*, 445.
[25] K. B. Wiberg, *Tetrahedron* **1968**, *24*, 1083.
[26] I. Mayer, *Chem. Phys. Lett.* **1983**, *270*, 97.
[27] M. Rafat, P. L. A. Popelier, In *The Quantum Theory of Atoms in Molecules. From Solid State to DNA and Drug Design*; C. F. Matta, R. J. Boyd, Eds.; WIley-VCH, Weinheim, Germany, **2007**; p. 121.
[28] A. Martín Pendás, E. Francisco, M. A. Blanco, *J. Comput. Chem.* **2005**, *26*, 344.
[29] E. Francisco, A. Costales, *Comput. Theor. Chem.* **2015**, *1053*, 77.
[30] I. Mayer, A. Hamza, *Int. J. Quantum Chem.* **2003**, *92*, 174.
[31] A. Martín Pendás, M. A. Blanco, E. Francisco, *J. Chem. Phys.* **2004**, *120*, 4581.
[32] C. D. H. Chisholm, *Group Theoretical Techniques in Quantum Chemistry;* Academic Press: London, **1976**. ISBN 0–12–172950–8.
[33] P. L. A. Popelier, D. S. Kosov, *J. Chem. Phys.* **2001**, *114*, 6539.
[34] P. L. A. Popelier, D. S. Kosov, *J. Chem. Phys.* **2000**, *113*, 3969.
[35] P. L. A. Popelier, L. Jouber, D. S. Kosov, *J. Phys. Chem. A* **2001**, *105*, 8254.
[36] A. Martín Pendás, E. Francisco, A QTAIM/IQA code (Available from the authors upon request by writing to ampendas@uniovi.es).
[37] C. Foroutan-Nejad, Z. Badri, R. Marek, *Phys. Chem. Chem. Phys.* **2015**, *17*, 30670.
[38] Z. Badri, C. Foroutan-Nejad, *Phys. Chem. Chem. Phys.* **2016**, *18*, 11693.
[39] M. Novák, C. Foroutan-Nejad, R. Marek, *J. Chem. Theory Comput.* **2016**, *12*, 3788.

Received: 11 November 2016
Revised: 24 January 2017
Accepted: 25 January 2017
Published online on 00 Month 2017

View Article Online
DOI: 10.1039/C6CP07617C

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

An unexpected bridge between chemical bonding indicators and electrical conductivity through the localization tensor†

Ángel Martín Pendás*, José Manuel Guevara-Vela, Daniel Menéndez Crespo, Aurora Costales and Evelio Francisco

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

While the modern theory of the insulating state shows that the conducting or insulating properties of a system can be extracted solely from ground state properties via the so-called localization tensor (LT), no chemical reading of this important quantity has ever been offered. Here, a remarkable link between the LT and bond orders as described by the delocalization indices (DIs) of chemical bonding theory is reported. This is achieved through a real space partition of the LT into intra- and interatomic contributions. We show that the convergence or divergence of the LT in the thermodynamic limit, which signals the insulating or conducting nature of an extended system, respectively, can be nailed down to DIs. This allows for the exploitation of traditional chemical intuition to identify essential and spectator atomic groups in determining electrical conductivity. The thermodynamic limit of the LT is controlled by the spatial decay rate of the interatomic DIs, exponential in insulators and power-law in conductors. Computational data in a few selected toy systems corroborate our results.

1 Introduction

As the technological demand of smart, functional, or taylored materials increases, so does the need for understanding the basic physics behind their sought-after properties. In many cases this search has led to explore the new dimension that the dependence of physical properties on size introduces at the nanoscale. For instance, the predicted demise of Moore's law[1] has stirred up the development of new quantum-mechanically operated devices like the single electron transistor.[2] Similarly, new fields such as molecular electronics have become hot topics producing thousands of specialized papers.[3] Despite much work, the building of new physical or chemical intuition that may guide future research beyond that coming from brute force case-by-case simulation has proven much more difficult. In crystal engineering, as an example, although the situation is now much better than 20 years ago, we are still far from mastering the rules to synthesize on-demand crystal structures.[4]

Regarding electrical conductivity at the nanoscale, much work has been devoted in molecular electronics to quantitatively simulate electron transport in single-molecule junctions,[5] and some rules regarding the factors that govern their conductivity have emerged. However, despite the efforts, no simple chemical rules linking molecular structure and molecular conductivity have been found todate.[3] Since, in the end, all newly developed nanodevices depend on the chemical synthesis of taylored molecular fragments, we believe that finding simple chemical indicators of facile electronic transport or conductivity is an important goal with possibly major outcomes.

A guiding principle in this quest may be taken from the naïve chemical association between conductivity and electron localization and delocalization. Key concepts in chemistry like conjugation, resonance, aromaticity, etc, are nothing but different incarnations of electron localizability. However, standard approaches coming from the theory of chemical bonding (TCB) are almost inevitably linked to the one-particle molecular orbital (MO) theory,[6] and molecular conductivity tends to be interpreted in terms of excitation gaps, i.e. HOMO-LUMO energetic differences, instead of as a ground state property that could be transformed into the sought conductivity indicators. Fortunately, a new paradigm in TCB has emerged in the last few decades[7] that defines (and explores) chemical objects in real space from orbital invariant densities (or density matrices). These techniques, collectively known as quantum chemical topology (QCT),[8] analyze the wave function of a system, and use meaningful fields to partition the

Departamento de Química Física y Analítica, Universidad de Oviedo, Oviedo. Spain. E-mail: ampendas@uniovi.es

† Electronic Supplementary Information (ESI) available: Mathematical derivations leading to the TPS in real space and details on its computational implementation in the PROMOLDEN code. See DOI: 10.1039/b000000x/

View Article Online
DOI: 10.1039/C6CP07617C

physical space into regions or domains associated to: atoms, through the one particle density in the quantum theory of atoms in molecules[7] (QTAIM); cores, lone, and bonding pairs, through the electron localization function[9] (ELF) or the electron localizability indicator[10] (ELI), etc. Once the real space objects are defined, indicators are obtained at well-defined points, usually the critical points of the defining field, and the global expectation values of operators are divided into domain contributions. This allows, for instance, for a rigorous real space partitioning of the energy into intra- and interatomic components (the interacting quantum atoms approach,[11] (IQA) much in the spirit of the atomistic ansatz. QCT, well-known in quantum chemistry, is slowly entering condensed matter physics.

Early attempts that tried to link the conductivity features of a molecular system with its electron density failed.[12] This comes as no surprise, since conductivity leaves no simple scars on the density. Fortunately, QCT domain expectation values are based on physical observables, so QCT provides an open door to connect the physicist's and the chemist's intuitions, which tend to live in separated worlds. This is not easy to do in other TCB approaches. In this regard, a rigorous formalism coupling the insulating or conducting nature of an extended system with ground state properties exists.[13] Although not well known in the chemical literature, Kohn's theory of the insulating state does the job. It is electron (de)localization that explains conductivity, quantified by an object called the localization tensor (LT).

Thanks to QCT and its rigorous partitioning of quantum mechanical expectation values into atomic or functional group contributions, we find and explore here a remarkable bridge between the LT and the standard bond orders of chemistry, as defined in their real space manifestation known as delocalization indices (DIs). It is the rate at which bond orders decrease with distance that determines whether a system will or will not be conducting in the static thermodynamic limit. Since we can examine straightforwardly the behavior of DIs among atoms or functional groups in several dimensions, we expect our results to be useful in building new conductivity chemical rules.

2 The modern theory of the insulating state and the assessment of conductivity via the localization tensor

A seminal work by W. Kohn in 1964[13] showed for the first time how the insulating nature of a system could be understood as a consequence of electron localization in the ground state, and not only from the properties of its excitation spectrum. However important, this line of reasoning remained largely unexplored until the end of the 1990's, when Resta revisited and generalized it.[14–16] As emphasized by this author,[17] it is the organization of electrons in the ground state that renders a system insulating or conducting. A central object that quantifies Kohn's localization in an N electron system is the localization tensor, $\boldsymbol{\lambda}$ or LT. It is defined as the second cumulant moment, per electron, of the total electronic position operator $\hat{\boldsymbol{R}} = \sum_i^N \hat{\boldsymbol{r}}_i$.

$$\boldsymbol{\lambda} = \frac{1}{N}\left\{\langle\Psi|\hat{\boldsymbol{R}}\otimes\hat{\boldsymbol{R}}|\Psi\rangle - \langle\Psi|\hat{\boldsymbol{R}}|\Psi\rangle\otimes\langle\Psi|\hat{\boldsymbol{R}}|\Psi\rangle\right\}. \tag{1}$$

We will use in this work bold fonts to indicate vectors or tensors, depending on the context, and the $\otimes$ symbol for tensor or cartesian products. As an example, the cartesian components of the $\boldsymbol{r}\otimes\boldsymbol{r}$ tensor are $(\boldsymbol{r}\otimes\boldsymbol{r})_{\alpha\beta} = x_\alpha x_\beta$.

One of the most important results of Resta's reformulation lies in the link between the behavior of $\boldsymbol{\lambda}$ in the thermodynamic limit and electrical conductivity: the $\boldsymbol{\lambda}$ tensor, that measures the quadratic fluctuation of the polarization of the system, and that was initially used by Kudinov,[18] has a well-defined thermodynamic limit, diverging for conductors while remaining finite for insulators.

We will just provide, for consistency, a few ideas that may guide the informed reader about the origin of such a unique property. It stems from the fluctuation-dissipation theorem,[19] that allows to prove[17] first that

$$\boldsymbol{\lambda} = \frac{\hbar}{\pi e^2 N}\int_0^\infty d\omega\ \mathrm{Im}\boldsymbol{\alpha}(\omega), \tag{2}$$

where $\boldsymbol{\alpha}(\omega)$ is the frequency dependent linear polarizability tensor. From this, if periodic boundary conditions are imposed, it can also be proven that

$$\lambda_{\beta\gamma} = \delta_{\beta\gamma}\frac{\hbar V}{\pi e^2 N}\int_0^\infty d\omega\ \frac{\mathrm{Re}\sigma(\omega)}{\omega}, \tag{3}$$

σ being the frequency dependent electric conductivity. For conducting systems, with non-vanishing Reσ at zero frequency, the diagonal components of $\boldsymbol{\lambda}$ diverge. These diagonal values can also be understood as localization lengths,[19] and they are related to the optical gap E_g by $\lambda_{\alpha\alpha} \leq \hbar^2/(2m_e E_g)$.

Simple manipulations, already put forward by Resta,[17] allow to recast the LT in terms of the first order, $\rho(\boldsymbol{r}_1)$, and the second order, $\rho_2(\boldsymbol{r}_1,\boldsymbol{r}_2)$, spinless densities, which are also known as the electron density and the pair density, respectively:

$$\begin{aligned}\rho(\boldsymbol{r}_1) &= N\sum_{\sigma_i}\int d\boldsymbol{x}_2\cdots d\boldsymbol{x}_N\Psi^*\Psi,\\ \rho_2(\boldsymbol{r}_1,\boldsymbol{r}_2) &= N(N-1)\sum_{\sigma_i}\int d\boldsymbol{x}_3\cdots d\boldsymbol{x}_N\Psi^*\Psi.\end{aligned} \tag{4}$$

In the above expressions $\Psi = \Psi(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)$ and we sum over all the σ_i spin components of the $\boldsymbol{x}_i$ space-spin electron coordinates.

Using the exchange-correlation density, the part of the pair density containing the quantum mechanical effects due to the antisymmetry of the wavefunction, $\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2) - \rho_2(\boldsymbol{r}_1,\boldsymbol{r}_2)$, and defining the interparticle position vector $\boldsymbol{r}_{12} = \boldsymbol{r}_1 - \boldsymbol{r}_2$, $\boldsymbol{\lambda}$ may be written in an explicitly origin independent, symmetrical form. As we show in the ESI, after some algebraic manipulations,

$$\boldsymbol{\lambda} = \frac{1}{2N}\int d\boldsymbol{r}_1 d\boldsymbol{r}_2\ (\boldsymbol{r}_{12}\otimes\boldsymbol{r}_{12})\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2). \tag{5}$$

3 A bridge between conductivity and the theory of chemical bonding

Being Chemistry the science of the interactions among electrons (or atoms, made from them and nuclei), it is not surprising that

View Article Online
DOI: 10.1039/C6CP07617C

ρ_{xc}, that collects all non-classical behaviour in the pair-density, is emerging, slowly but steadily, as one of the pillars in the modern TCB. Similarly, having shown that electrical conductivity is related to how electrons localize or delocalize, it is also expectable that the former be related to chemical bonding measures of electron delocalization. Undoubtedly, the latter have a rather long history in TCB.

In the present context, Bader and Stephens[20] already proposed in 1974 that the interatomic integration of ρ_{xc} measures the number of pairs of electrons shared between two atomic regions, and named this quantity the delocalization index (DI),

$$\delta^{AB} = 2\int_A d\boldsymbol{r}_1 \int_B d\boldsymbol{r}_2\, \rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2). \tag{6}$$

Here A,B are the spatial regions associated to the two atoms (or fragments) under scrutiny. A similar A,A integral, the localization index, determines the number of localized electrons in a region. The DI provides the fluctuation of the electron population in the A,B regions, being also a real space generalization of the standard MO Wiberg-Mayer bond order,[21,21,22] which physicists still use, in their majority, in its even cruder Mulliken flavor. In energetic terms, the interatomic exchange-correlation energy,

$$E_{xc}^{AB} = \int_A d\boldsymbol{r}_1 \int_B d\boldsymbol{r}_2\, \frac{\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2)}{r_{12}} \tag{7}$$

has been shown to correspond to the covalent part of the interaction between the regions.[23] Besides these two direct links, a growing body of evidence is showing the relevance of ρ_{xc}-based indices in TCB,[24] explaining facts as the nature of chemical interactions from DI profiles,[24] or rationalizing stereolectronic effects.[25] As we are going to show, *it is the innocent concept of bond order, a must in every freshman chemistry course, that stores information about electrical conductivity*.

A couple of recent studies,[26,27] had already started to show that DIs encode information about the insulating or conducting nature of a system through their spatial decay rate: in metals we find an algebraic oscillatory decline with interatomic $A-B$ distance, while in insulators their fall off is exponential. Consideration of strongly correlated cases[27] evidences that DIs are also suitable generalizations of the double occupation order parameter $\mathscr{D}$ used in Hubbard models to signal metal-insulator transitions, and that they reveal how mesomeric effects in alternant hydrocarbons are deeply linked to the oscillatory pattern that leads to conductivity in the thermodynamic limit. An increase in the electron correlation strength (by increasing the Hubbard U/t parameter or equivalently by substituting carbon by heavier elements) eventually destroys the oscillations, pointing toward an active effect of electron correlation in chemistry, e.g. to smaller mesomeric effects in the heavier analogues of alternant hydrocarbons. Even more importantly, the decay of these indices may be followed along specific bond chains, directions, or along a combination of both. One needs only choose appropriately the domains in the A,B pairs. This provides a quantitative tool in the discovery of low dimensional conductors.

3.1 The localization tensor in finite molecules

In the last few years, several works by Leininger, Evangelisti and coworkers[28,29] have examined the role of $\boldsymbol{\lambda}$ in molecular instead of extended systems. To that end, these authors have preferred to use the total second cumulant, which they have called total position-spread tensor, $\boldsymbol{\Lambda}$ or TPS, and not the per electron quantity. Even a spin resolved version has also been studied.[30]

It has been shown that the TPS is very sensitive to bond stretching, becoming large in the case of increased electron mobility. In simple diatomics, for instance, its parallel component is small at equilibrium, it increases as the interatomic distance is enlarged before achieving a maximum value close to the bond breaking region, and it decreases again towards the free atomic value at dissociation.

Although the TPS has been wellcome, adding to the battery of new chemical bonding indicators at hand, its global character partially limits its applicability. Its evolution in a possibly complex process will just average out the total response of the system, even though some very restricted atomic or bond resolution might be achieved by following a particular component or projection that isolates an important direction in space. In order to become a useful TCB descriptor, this barrier needs be overcome to understand the origin of convergence/divergence and the onset of conductivity as we approach the thermodynamic limit.

3.2 An atomic partition of the TPS

Being the expectation value of a two-electron operator, QCT offers an immediate solution to the problem: provided that a chemically meaningful division of the space exists, we can space partition $\boldsymbol{\lambda}$ or $\boldsymbol{\Lambda}$, just as it is done in the IQA approach. Without loss of generality, we present an atomic partition of the TPS using the QTAIM. This can be made coarser (scaling it up to the functional group or molecular level) or finer (to the level of atomic core, bond and lone pair domains) at will. Another important point regards origin dependency, which may bring trouble in the partitioning if direct use of Eq. 1 is made. This difficulty is eluded by using the manifestly origin independent Eq. 5.

Let us start with an exhaustive partition of the physical space $R^3 = \bigcup A$ into atomic regions. A rigorous, physically sound possibility is provided by the QTAIM. Each of these regions or domains harbors a nucleus, at position $\boldsymbol{R}_A$. Given the one-to-one correspondence between domains and nuclei, we will label them interchangeably. Then,

$$\begin{aligned}
\boldsymbol{\Lambda} &= \sum_{A\geq B} \boldsymbol{\Lambda}^{AB}, \qquad (8)\\
\boldsymbol{\Lambda}^{AA} &= \frac{1}{2}\int_A d\boldsymbol{r}_1 \int_A d\boldsymbol{r}_2\, (\boldsymbol{r}_{12}\otimes\boldsymbol{r}_{12})\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2),\\
\boldsymbol{\Lambda}^{AB} &= \int_A d\boldsymbol{r}_1 \int_B d\boldsymbol{r}_2\, (\boldsymbol{r}_{12}\otimes\boldsymbol{r}_{12})\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2).
\end{aligned}$$

Notice that the above expressions provide a chemical partition of the TPS (or the LT if we divide by N).

The intra-atomic $\boldsymbol{\Lambda}^{AA}$ terms must tend to their free atomic values $\boldsymbol{\Lambda}_0^{AA}$ as the molecular system is pulled apart into atoms.

View Article Online
DOI: 10.1039/C6CP07617C

Recalling that $\int d\boldsymbol{r}_1 d\boldsymbol{r}_2 \rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2) = N$, it is well known that the localization index of region A, $N_{AA} = \int_A d\boldsymbol{r}_1 \int_A d\boldsymbol{r}_2 \rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2)$, defines the number of localized electrons in region A, so that $\boldsymbol{\Lambda}^{AA}$ measures the interelectron spread of these localized electrons in the atomic region, behaving grossly as $\boldsymbol{\Lambda}^{AA} \sim (N_{AA}/2)\langle r_{12}^2\rangle_A$. The intra-atomic contributions to the TPS are thus additive and size extensive, and their sum is clearly seen through this partition to provide a term that scales linearly with the size of the system (or the number of electrons) as we approach the thermodynamic limit. An important corollary is that the root of any divergence in the LT will not be found in these intra-atomic components (see below). As it happens with other intra-domain expectation values in QCT, $\Delta\boldsymbol{\Lambda}^{AA} = \boldsymbol{\Lambda}^{AA} - \boldsymbol{\Lambda}_0^{AA}$ reflects the local change in interelectron spread due to chemical bonding and, except in very specific cases, like those in which a large charge transfer occurs, we expect these $\Delta\boldsymbol{\Lambda}^{AA}$'s to be small.

The interatomic $\boldsymbol{\Lambda}^{AB}$ contributions are much more interesting, for they directly measure the change in the interelectron spread due to the delocalization associated to the formation (or breaking) of a particular bond. From chemical intuition, two spatially separated non-bonded atoms will display a vanishing $\boldsymbol{\Lambda}^{AB}$ value. Most, if not all, of the interesting behavior of $\boldsymbol{\Lambda}$ is then to be found in these terms. Notice that when two separated entities interact the Λ^{AA} components are non-vanishing at full-separation, changing upon interaction, but that the Λ^{AB} value is zero at infinite separation, sensing directly the interaction process.

3.3 How does $\boldsymbol{\Lambda}^{AB}$ decay with distance?

Taking into account that A and B are two non-overlapping regions of space, the behavior of $\boldsymbol{\Lambda}^{AB}$ as the two centers move away from each other is determined by the decay rate of ρ_{xc} with interelectron distance and the $\boldsymbol{R}_{AB} = \boldsymbol{R}_A - \boldsymbol{R}_B$ distance itself. Provided that the two electron coordinates satisfy $\boldsymbol{r}_1 \in A$ and $\boldsymbol{r}_2 \in B$, we may refer them to their local nuclear reference frames, respectively: $\boldsymbol{r}_1 = \boldsymbol{R}_A + \boldsymbol{u}_1$, $\boldsymbol{r}_2 = \boldsymbol{R}_B + \boldsymbol{u}_2$. With the above, the dependency of $\boldsymbol{\Lambda}^{AB}$ on the internuclear distance is explicitly separated. Let us define $\boldsymbol{u}_{12} = \boldsymbol{u}_1 - \boldsymbol{u}_2$, and the local integrals

$$\begin{aligned} \boldsymbol{I} &= \int_A d\boldsymbol{u}_1 \int_B d\boldsymbol{u}_2\ (\boldsymbol{u}_{12}\otimes\boldsymbol{u}_{12})\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2), \\ \boldsymbol{J} &= \int_A d\boldsymbol{u}_1 \int_B d\boldsymbol{u}_2\ \boldsymbol{u}_{12}\,\rho_{xc}(\boldsymbol{r}_1,\boldsymbol{r}_2), \end{aligned} \tag{9}$$

that may also be written in terms of spatial moments of the domain averaged Fermi holes introduced by R. Ponec,[31] which have been successfully used in the last years to reveal many interesting effects in chemical bonding.[32] With these, we may write

$$\boldsymbol{\Lambda}^{AB} = \boldsymbol{I} + \boldsymbol{R}_{AB}\otimes\boldsymbol{J} + \boldsymbol{J}\otimes\boldsymbol{R}_{AB} + \frac{1}{2}(\boldsymbol{R}_{AB}\otimes\boldsymbol{R}_{AB})\delta^{AB}. \tag{10}$$

The first term contains only local distances, roughly decaying as δ^{AB} itself, and out of the three remaining terms, the one leading the long-range behavior is the third. Thus, at large interatomic distances $\boldsymbol{\Lambda}^{AB} \sim (\boldsymbol{R}_{AB}\otimes\boldsymbol{R}_{AB})\delta^{AB}/2$, and the parallel component of $\boldsymbol{\Lambda}$ along the bond direction will scale as

$$\Lambda_{\parallel}^{AB} \sim \frac{1}{2}R_{AB}^2\delta^{AB}. \tag{11}$$

This last important relation provides a new bridge between TCB descriptors in the ground state and the Kohn-Resta theory of the insulating state.

3.4 The chemical bonding origin of the convergence/divergence of $\boldsymbol{\lambda}$

We can now turn to the convergence/divergence of $\boldsymbol{\lambda}$ in the thermodynamic limit. An explicit effective one-center expansion of Λ may be immediately written from our previous partition as

$$\boldsymbol{\Lambda} = \sum_A \boldsymbol{\Lambda}^A,\ \boldsymbol{\Lambda}^A = \boldsymbol{\Lambda}^{AA} + \frac{1}{2}\sum_{B\neq A}\boldsymbol{\Lambda}^{AB}. \tag{12}$$

Notice that the atomic additivity of $\boldsymbol{\Lambda}^A$ allows us to write

$$\boldsymbol{\lambda} = \frac{1}{N}\sum_A^{N_{at}}\boldsymbol{\Lambda}^A = \frac{N_{at}}{N}\langle\boldsymbol{\Lambda}^A\rangle = \frac{\langle\boldsymbol{\Lambda}^A\rangle}{n}, \tag{13}$$

where $\langle\boldsymbol{\Lambda}^A\rangle$ is the average of $\boldsymbol{\Lambda}^A$ over all the atoms comprising our system, and n is the average number of electrons per atom, The divergence of $\boldsymbol{\lambda}$ in the thermodynamic limit is equivalent to that of the average atomic-additive $\langle\boldsymbol{\Lambda}^A\rangle$. Remarkably, the divergence of the LT can thus be nailed down to an atomic property.

Further analysis opens new avenues in understanding the onset of conductivity from a chemical perspective. Several paths may lead to a divergent $\langle\boldsymbol{\Lambda}^A\rangle$. For instance, all of the $\boldsymbol{\Lambda}^A$ terms may diverge themselves, or only one or a few. This analysis will identify *essential and spectator* atoms or functional groups in complex conducting systems. Essential groups for conductivity will be those for which $\boldsymbol{\Lambda}^A$ diverges, while spectator groups will be characterized by convergent $\boldsymbol{\Lambda}^A$. We think that this classification scheme can help identify replaceable groups that will not change the basic conductivity properties of a system while tuning their fine conductive properties.

For each divergent $\boldsymbol{\Lambda}^A$, our previous comments show that it will be the interatomic sum, $\sum_{B\neq A}\boldsymbol{\Lambda}^{AB}$, not the intra-atomic $\boldsymbol{\Lambda}^{AA}$, that will add to an infinite result. It is the interplay between the dimensionality of the system and the decay rate of δ^{AB}, that determines convergence. This binds the behavior of $\boldsymbol{\lambda}$ to the decay rate of DIs, already explored.[26,27]

To keep our discussion as simple as possible, we will now continue our reasoning in one-dimensional systems, where these ideas are most easily aprehended. In 1D, whenever δ^{AB} decreases faster than $\delta^{AB} \simeq R_{AB}^{-d}$, with $d = 2$, then the $\sum_{B\neq A}\boldsymbol{\Lambda}^{AB}$ term will converge, and the contrary will make it diverge. Similarly, the limiting d exponent is 3,4 for 2- and 3-dimensional conductivity to occur, respectively. These results perfectly match the findings relating the decay rate of the non-diagonal elements of the first order density in tight binding models of metals, as shown by Taraskin.[33,34] On the contrary, exponentially decaying interatomic delocalization indices δ^{AB} will always lead to insulating behavior, i.e. to convergent $\boldsymbol{\lambda}$ values. We would like to stress that

View Article Online
DOI: 10.1039/C6CP07617C

the transition from exponential to power-law δ^{AB} decay rates has already been found to occur in computational studies of model systems.[27]

The following expressions summarize the core of our findings in the thermodynamic limit:

- If δ^{AB} decays exponentially with R^{AB}, $\boldsymbol{\lambda}$ converges.
- In a d-dimensional system, $\boldsymbol{\lambda}$ converges/diverges if δ^{AB} decays faster/slower than $R_{AB}^{-(d+1)}$.

We propose that low-dimensional, as well as bulk conductivity can be spotted by examining the behavior of $\boldsymbol{\Lambda}^{AB}$ along the appropriate directions, planes, or 3D regions, respectively.

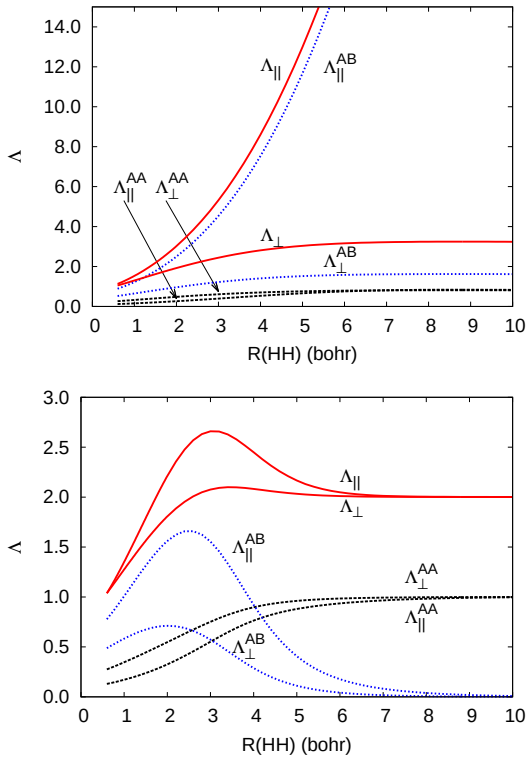


Fig. 1 Total, intra-atomic, and interatomic components of $\boldsymbol{\Lambda}$ in the H_2 molecule at the HF (top) and CASSCF//aug-cc-pVTZ (bottom) levels along the internuclear dissociation coordinate. The parallel and perpendicular labels correspond to the internuclear and orthogonal directions, respectively. All data in au.

4 Exemplifying the usefulness of the Λ partition

We are now in a position to show the new insights at work in a few toy systems. We will first discuss two simple dissociation processes, $H_2 \longrightarrow H + H$ and $H_2O \longrightarrow OH + H$, where we will see how, and why, the final localization of electrons in the products leads to a convergent $\boldsymbol{\lambda}$, i.e. to insulating-like behavior. This will also make the essential role of electron correlation in correctly capturing the physics of the system. Finally, the early stages in the birth of a divergence, thus the switch towards metallic-like behavior, will be succintcly analyzed in a linear chain of equally spaced hydrogen atoms. The electronic structure calculations have been performed with the GAMESS package,[35] and the TPSs have been obtained for QTAIM atomic partitions through our PROMOLDEN,[36] code, which is able to handle quite a number of correlated and non-correlated wave functions and several QCT partitions, not only the one provided by the QTAIM. Details of the implementation of the TPS in PROMOLDEN can be found in the ESI.

4.1 The dissociation of H_2

First we discuss the H_2 molecule ($A = H$, $B = H'$), a paradigm of covalent interactions. We have computed $\boldsymbol{\Lambda}$ at the Hartree-Fock (HF) and the configuration active space (CASSCF) levels with the aug-cc-pVTZ basis set along its dissociation coordinate. The results are contained in Fig. 1. Notice that the $\boldsymbol{\Lambda}$ tensor is diagonal in any reference frame in which the internuclear distance coincides with one of the coordinate axes, and that rotational invariance equalizes the other two orthogonal eigenvalues of $\boldsymbol{\Lambda}$. We will call these two different components of $\boldsymbol{\Lambda}$ $\Lambda_{||}$ and $\Lambda_{\perp}$, respectively. As already put forward by Resta[17] and Leininger *et. al.*,[28] mean-field and correlated descriptions of the dissociation process differ essentially. Interestingly, this qualitative differences are also observed when the delocalization index is examined by itself.[37].

Failure to consider electron correlation leads to a parabolic divergence of $\Lambda_{||}$ as the internuclear distance increases. Its origin cannot be grasped by solely examining the full tensor, but its partitioning shows that, as expected, it is the interatomic component, $\Lambda_{||}^{AB}$, that diverges. Through the eyes of our findings, the parabolic behavior is due to an artificial non-vanishing δ^{AB} at infinite separation (the HF dissociation error). It may be instructive to recall that at the Hartree-Fock level we may write for closed-shell systems

$$\delta^{AB} = 4\sum_{ij} S_{ij}^{A} S_{ij}^{B}, \tag{14}$$

where the sum runs over all pairs ij of occupied orbitals and S_{ij}^{A} is an atomic overlap integral,

$$S_{ij}^{A} = \int_{A} d\boldsymbol{r}\, \phi_i(\boldsymbol{r})\phi_j(\boldsymbol{r}). \tag{15}$$

As in H_2 we have only one occupied HF orbital fulfilling $S_{11}^{A} = S_{11}^{B} = 1/2$ by symmetry considerations, the bond order $\delta^{AB} = 1$ at any R_{HH} in this model. This can be also interpreted as the result of the two opposite spin electrons being statistically independent if no Coulomb correlation is added. The wrong constant δ^{AB} leads to infinite-range delocalization, with an overall probability of finding the two electrons in any one of the H atoms (the so-called ionic weight in quantum chemical approaches) equal to $1/2$.[38] Eq. 11 does the rest.

Proper inclusion of Coulomb correlation makes δ^{AB} decrease exponentially at large distances,[39] so that the bond breaks ap-

View Article Online
DOI: 10.1039/C6CP07617C

propriately. Thus, in the correct correlated description, the intra-atomic $\mathbf{\Lambda}^{AA}$ components start at low values close to equilibrium, increasing to the free atom limit. From Eq. 1, it is clear that in this limit $\mathbf{\Lambda}^{AA}$ is also diagonal and that each of its three components is equal to $\langle\phi|r^2|\phi\rangle$, where ϕ is the hydrogenic atomic orbital in the state of interest. For the 1s ground state, $\langle r^2\rangle = 3$ au. so that $\langle z^2\rangle = 1$, which is also the appropriate $\mathbf{\Lambda}^{AA}$ limit at dissociation.

Many other features of the behavior of the intra-atomic components are easy to rationalize. For instance, the lower value of $\mathbf{\Lambda}^{AA}$ at small interatomic distances is understood straightforwardly, for the number of localized electrons in each atom is in these conditions about 0.5, *vide supra*. It is also interesting to notice that the intra-atomic $\Lambda_{\parallel}^{AA}$ value is smaller than its $\Lambda_{\perp}^{AA}$ counterpart, this reflecting the compression of the atomic density along the internuclear axis as we approach the two atoms from infinity. As also expected, it is the interatomic $\mathbf{\Lambda}^{AB}$ component that accounts for the sharp maximum in the total $\mathbf{\Lambda}$ tensor. This maximum has been interpreted[28] as a signature of bond breaking. Under our present formalism, it is a simple consequence of the shift from a power-law to an exponential decay in ρ_{xc} or δ^{AB}, i.e. from a quasi-independent electron pair being stretched (as in the HF case) to the strongly correlated, localized dissociation limit. This transition, scaled by $R_{AB}^2/2$, gives rise to the maximum, that is found very close to the internuclear distance at which the inflection point of δ^{AB} has been repeatedly described.[24,27]

4.2 The power of partitioning Λ: $H_2O \longrightarrow OH + H$

We can now show how the global behavior of $\mathbf{\Lambda}$, that contains the total response of the system to a chemical process, may be split up into chemically meaningful terms. To that end we have chosen the $H_2O \longrightarrow OH + H$ dissociation, with the OH1 distance taken as an intrinsic reaction coordinate. Fig. 2 shows the evolution of $\Lambda_{\parallel}^{AB}$ along R(OH1) at the complete active space CASSCF[8,8]//aug-cc-pVTZ level. This exemplifies the power of partitioning Λ. As the OH1 distance is stretched, we see how there is a simple jump in $\Lambda_{\parallel}^{OH2}$, while it is $\Lambda_{\parallel}^{OH1}$ that behaves much as in the H_2 case. The step from lower to higher $\Lambda_{\parallel}^{OH2}$ can be understood by taking into account that in the final OH radical the number of delocalized electrons between the O and the H2 atoms has increased. In other words, since the OH2 bond order increases as the H1 atom dissociates, $\Delta\delta^{OH2} > 0$, so does $\Lambda_{\parallel}^{OH2}$. Only a partitioning of the TPS, like the one devised here, will be able to isolate the main actors in complex scenarios. With our tools, this seems to be at hand, and the strong link between the essential interatomic $\mathbf{\Lambda}^{AB}$ terms and the DIs is unveiled.

4.3 Recognizing the onset of conductivity: The H_{10} chain

Our next example will be a linear chain of 10 equally spaced H atoms computed at the HF and full valence CASSCF levels with the 6-311G* basis set. At the inter-hydrogen distance selected, $R = 3.5$ bohr, the HF model is starting to fail, but it still provides a reasonable description of the electron system. We examine how the interlectron spread propagates along a quasi-1D system, and our goal is put on Eq. 12. Fig 3 shows how $\Lambda_{\parallel}^{AB}$ changes for all pairs in which one of the atoms is fixed to be an end H. The first

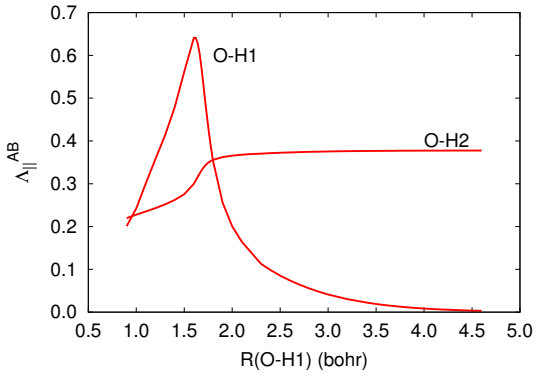


Fig. 2 Interatomic components of $\Lambda_{\parallel}$ in the H_2O molecule at the CASSCF[8,8]//aug-cc-pVTZ level along the internuclear O-H_1 dissociation coordinate. All data in au.

interesting point is that in the mean-field HF approximation, $\Lambda_{\parallel}^{AB}$ decays with distance in a well-developed slow oscillatory pattern. We have found a similar behavior examining DIs.[26,27]

Notice that, from the chemical point of view, these oscillations signal a clear bond order alternation or mesomerism, precursor of a Peierls distortion (or H_2 dimerization). This has also been repeatedly described in previous literature:[24] delocalization indices in geometrically constrained systems inform about the expected distortions when the constraints are released. As we find here, the localization tensor yields similar sensible chemical information.

A power law fit of δ^{AB} to R_{AB}^{-d} gives $d \approx 2.5$, close to the tight binding value ($d = 2.0$). The sum in Eq. 12 achieves a very large value. The oscillatory pattern in $\Lambda_{\parallel}^{AB}$, as in the case of the DI, is a clear indicator of conducting-like behavior. A second point regards the very quick saturation of $\mathbf{\Lambda}^{AB}$ to the $R_{AB}^2\delta^{AB}/2$ leading term in the long range. Fig. 3 shows that our previous theoretical insights are fully realized from actual computations. It is the decay rate of DIs (i.e. the inter-center electron delocalization) that determines conductivity in the thermodynamic limit. Inclusion of electron correlation does not make the oscillations disappear at this interatomic distance, but reveals how the electrons are now much more localized, with a considerably smaller spread. A similar fit now gives $d \approx 4.1$, well above the metallic limit. We have shown that the oscillations disappear when we enter the dissociating, localized regime, and that the DI decays exponentially in that case.[27]

4.4 Insulator-like and conducting-like chains

We will finish our discussion by considering two real life one-dimensional linear chains of equidistant atoms: $(LiH)_{15}$ and Li_{10}, with nearest neighbor distances set to 3.0 and 5.818 bohr, respectively. We have used a HF/6-311G* level that provides a simple, yet reasonable description of both systems. Fig. 4 shows relevant values for $\Lambda_{\parallel}^{AB}$, that provide clear grounds for comparison: the small values and the very quick decay of the interatomic $\Lambda_{\parallel}^{AB}$ values in an insulator like lithium hydride, and their much larger

View Article Online
DOI: 10.1039/C6CP07617C

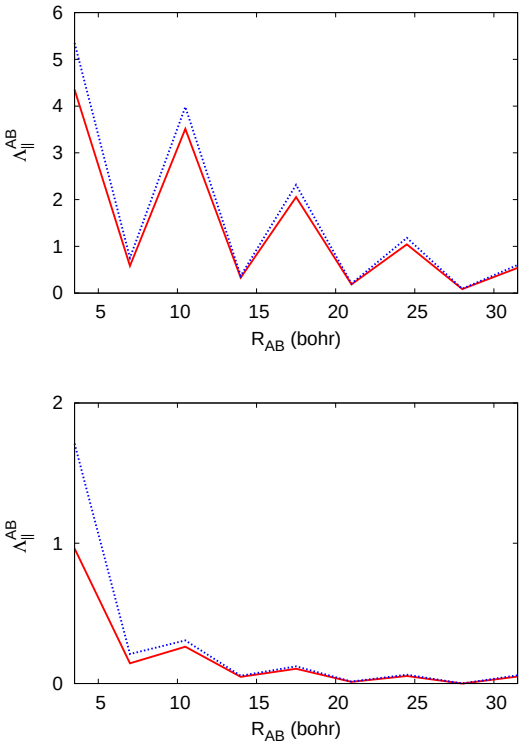


Fig. 3 $\Lambda_{\parallel}^{AB}$ along the internuclear direction in a H_{10} linear chain of equally space atoms at the HF//6-311G* (top) and CASSCF//6-311G* (bottom) levels. A labels one of the end H atoms, and B runs over all the others. Nearest neighbohrs are 3.5 bohr apart. In each plot, the dashed curve corresponds to $R_{AB}^2\delta^{AB}/2$. All data in au.

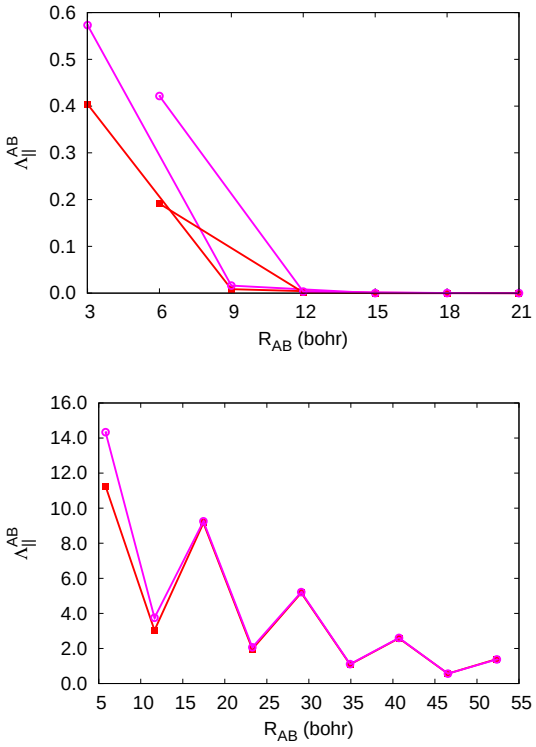


Fig. 4 HF//6-311G* $\Lambda_{\parallel}^{AB}$ (bold squares) and $R_{AB}^2\delta^{AB}/2$ (circles) along the internuclear direction. Top: linear $(LiH)_{15}$ chain of equally spaced nearest-neighbors atoms. Values starting at $R_{AB} = 3.0/6.0$ bohr correspond to the AB=LiH/HH pairs, respectively. Bottom: linear Li_{10} chain of equidistant Li atoms separated by 5.818 bohr. The AB pairs are those formed by one end Li atom and all of its neighbors.

magnitude, slow decay, and oscillatory behavior in the metallic-like chain.

5 Conclusions

Summarizing, we have shown in this article that a remarkable bridge exists between the Kohn-Resta theory of the insulating state, through the localization tensor (or its total position spread tensor version in molecular systems), and well known indicators used in the modern theory of chemical bonding as bond orders. This has been achieved by partitioning the localization tensor in intra- and interatomic components. An orbital invariant way to do so starts by writing the LT in terms of reduced densities and then partitioning the space into atomic regions according to quantum chemical topology. Convergence or divergence of the LT in the thermodynamic limit , associated to insulating or conducting electrical properties, depends exclusively on the decay rate of its interatomic components. The latter are dominated by the chemical delocalization index, a modern form of bond order. The chemistry of ground states and the physics of conductivity become intertwined in this way. We expect this new link to be useful in the search and design of low dimension conductors or insulators, for the total LT can be written as a sum of atomic (or functional group) components. Each atom or functional group in a system may thus be classified as essential, if its contribution to the LT diverges, or an spectator, if it converges, as electrical conductivity is regarded. We expect that this categorization can be used advantageously in the rational design of new materials.

Acknowledgements: We thank the spanish MINECO/FEDER, grant CTQ2015-65790-P, FICyt, grant GruPin14-049, and Conacyt/Mexico, grant 381483, for financial support.

References

1 M. M. Waldrop, *Nature*, 2016, **530**, 144.
2 T. Fulton and G. Dolan, *Phys. Rev. Lett.*, 1987, **109**, 59.
3 J. C. Cuevas and E. Scheer, *Molecular Electronics. An Introduction to Theory and Experiment*, World Scientific, Singapore, 2010.
4 G. R. Desirajou, J. J. Vittal and A. Ramanan, *Crystal Engineering: A Textbook*, World Scientific, Singapore, 2011.

View Article Online
DOI: 10.1039/C6CP07617C

5 L. Venkataraman, J. E. Klare, C. Nuckolls, M. S. Hybertsen and M. L. Steigerwald, *Nature*, 2006, **422**, 904.
6 B. M. Gimarc, *Molecular structure and bonding. The qualitative molecular orbital approach*, Academic Press, New York, 1979.
7 R. F. W. Bader, *Atoms in Molecules*, Oxford University Press, Oxford, 1990.
8 P. L. A. Popelier and E. A. Brèmond, *Int. J. Quant. Chem.*, 2009, **109**, 2542.
9 B. Silvi and A. Savin, *Nature*, 1994, **371**, 683.
10 M. Kohout, *Int. J. Quant. Chem.*, 2004, **97**, 651.
11 M. A. Blanco, A. Martín Pendás and E. Francisco, *J. Chem. Theory Comput.*, 2005, **1**, 1096.
12 W. Cao, C. Gatti, P. MacDougall and R. Bader, *Chem. Phys. Lett.*, 1987, **141**, 380.
13 W. Kohn, *Phys. Rev.*, 1964, **133**, A171.
14 R. Resta, *Phys. Rev. Lett.*, 1998, **80**, 1800.
15 R. Resta and S. Sorella, *Phys. Rev. Lett.*, 1999, **82**, 370.
16 R. Resta, *Phys. Rev. Lett.*, 2005, **95**, 196805.
17 R. Resta, *J. Chem. Phys.*, 2006, **124**, 104104.
18 E. K. Kudinov, *Sov. Phys. Solid State*, 1991, **33**, 1299.
19 I. Souza, T. Wilkens and R. M. Martin, *Phys. Rev. B*, 2000, **62**, 1666.
20 R. F. W. Bader and M. E. Stephens, *Chem. Phys. Lett.*, 1974, **26**, 445.
21 K. B. Wiberg, *Tetrahedron*, 1968, **24**, 1083.
22 I. Mayer, *J. Comput. Chem.*, 2007, **28**, 204–221.
23 A. Martín Pendás, E. Francisco and M. A. Blanco, *J. Phys. Chem. A*, 2006, **110**, 12864.
24 M. García-Revilla, P. L. A. Popelier, E. Francisco and A. Martín Pendás, *J. Chem. Theory Comput.*, 2011, **7**, 1704.
25 A. Martín Pendás, M. A. Blanco and E. Francisco, *J. Comput. Chem.*, 2009, **30**, 98.
26 A. Gallo-Bueno and A. Martín Pendás, *Phys. Chem. Chem. Phys.*, 2016, **18**, 11772.
27 A. Gallo-Bueno, M. Kohout and A. Martín Pendás, *J. Chem. Theory Comput.*, 2016, **12**, 3053.
28 O. Brea, M. El Khatib, C. Angeli, G. L. Bendazolli, S. Evangelisti and T. Leininger, *J. Chem. Theory Comput.*, 2013, **9**, 5286.
29 M. El Khatib, T. Leininger, G. L. Bendazolli and S. Evangelisti, *Chem. Phys. Lett.*, 2014, **591**, 58.
30 M. El Khatib, O. Brea, E. Fertitta, G. L. Bendazzoli, S. Evangelisti and T. Leininger, *J. Chem. Phys.*, 2015, **142**, 094113.
31 R. Ponec, *J. Math. Chem.*, 1997, **21**, 323.
32 E. Francisco, A. Martín Pendás and M. A. Blanco, *J. Chem. Phys.*, 2009, **131**, 124125.
33 S. N. Taraskin, P. A. Fry, X. Zhang, D. A. Drabold and S. R. Elliot, *Phys. Rev. B*, 2002, **66**, 233101.
34 S. N. Taraskin, D. A. Drabold and S. R. Elliot, *Phys. Rev. Lett.*, 2002, **88**, 196405.
35 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery, *J. Comput. Chem.*, 1993, **14**, 1347.
36 A. Martín Pendás and E. Francisco, A QTAIM/IQA code (Available from the authors upon request).
37 A. Martín Pendás, E. Francisco and M. A. Blanco, *Chem. Phys. Lett.*, 2007, **437**, 287.
38 E. Francisco, A. Martín Pendás and M. A. Blanco, *J. Chem. Phys.*, 2007, **126**, 094102.
39 A. Martín Pendás, E. Francisco and M. A. Blanco, *Faraday Discuss.*, 2007, **135**, 423.

Part VI

Bibliography

BIBLIOGRAPHY

[1] Alcoba, D. R., Lain, L., Torre, A. and Bochicchio, R. C. (2009), «Local spin: A treatment beyond single determinant wave functions», *Chem. Phys. Lett.*, vol. 470 (1-3), pp. 136–139, URL http://dx.doi.org/10.1016/j.cplett.2009.01.034http://linkinghub.elsevier.com/retrieve/pii/S000926140900044X. {Cited on page 100.}

[2] Alcoba, D. R., Torre, A., Lain, L. and Bochicchio, R. C. (2011), «Descriptions of local spins in the three-dimensional physical space», *Chem. Phys. Lett.*, vol. 504 (4-6), pp. 236–240, URL http://dx.doi.org/10.1016/j.cplett.2011.01.071http://linkinghub.elsevier.com/retrieve/pii/S0009261411001138. {Cited on page 100.}

[3] Alves, H., Molinari, A. S., Xie, H. and Morpurgo, A. F. (2008), «Metallic conduction at organic charge-transfer interfaces», *Nat Mater*, vol. 7 (7), pp. 574–580, URL http://www.ncbi.nlm.nih.gov/pubmed/18552852. {Cited on page 182.}

[4] Ángyán, J. G., Gerber, I. and Marsman, M. (2006), «Spherical harmonic expansion of short-range screened Coulomb interactions», *J. Phys. A. Math. Gen.*, vol. 39 (27), pp. 8613–8630, URL http://stacks.iop.org/0305-4470/39/i=27/a=005?key=crossref.990ab9897fb3bf8c1276da35185de4a5. {Cited on page 55.}

[5] Apeloig, Y., Pauncz, R., Karni, M., West, R., Steiner, W. and Chapman, D. (2003), «Why Is Methylene a Ground State Triplet while Silylene Is a Ground State Singlet? †», *Organometallics*, vol. 22 (16), pp. 3250–3256, URL http://pubs.acs.org/doi/abs/10.1021/om0302591. {Cited on page 112.}

[6] Aslangul, C., Daudel, R., Constanciel, R. and Kottis, P. (1972), «Aspects of the Localizability of Electrons in Atoms and Molecules: Loge Theory and Related Methods», *Adv. Quantum Chem.*, vol. 6, pp. 93–141. {Cited on page 77.}

[7] Bader, R. (1994), *Atoms in Molecules: A Quantum Theory*, International Ser. of Monogr. on Chem, Oxford University Press, Incorporated.

[8] Bader, R. and Stephens, M. (1974), «Fluctuation and correlation of electrons in molecular systems», *Chem. Phys. Lett.*, vol. 26 (3), pp. 445–449, URL http://linkinghub.elsevier.com/retrieve/pii/000926147489069X. {Cited on page 70.}

[9] Bader, R. and Stephens, M. (1974), «Fluctuation and correlation of electrons in molecular systems», *Chem. Phys. Lett.*, vol. 26 (3), pp. 445–449, URL http://linkinghub.elsevier.com/retrieve/pii/000926147489069X. {Cited on pages 121 and 149.}

[10] Bader, R. F. (1990), *Atoms in molecules. A Quantum Theory*, Clarendon Press. {Cited on pages 120 and 147.}

[11] Bader, R. F. W. (1991), «A quantum theory of molecular structure and its applications», *Chem. Rev.*, vol. 91 (5), pp. 893–928, URL http://pubs.acs.org/doi/abs/10.1021/cr00005a013. {Cited on page 78.}

[12] Bader, R. F. W. and Beddall, P. M. (1972), «Virial Field Relationship for Molecular Charge Distributions and the Spatial Partitioning of Molecular Properties», *J. Chem. Phys.*, vol. 56 (7), pp. 3320–3329, URL http://link.aip.org/link/?JCP/56/3320/1{&}Agg=doihttp://aip.scitation.org/doi/10.1063/1.1677699. {Cited on page 78.}

[13] Badri, Z. and Foroutan-Nejad, C. (2016), «Unification of ground-state aromaticity criteria – structure, electron delocalization, and energy – in light of the quantum chemical topology», *Phys. Chem. Chem. Phys.*, vol. 18 (17), pp. 11 693–11 699, URL http://xlink.rsc.org/?DOI=C5CP05222J. {Cited on page 141.}

[14] Baerends, E. J. and Gritsenko, O. V. (1997), «A Quantum Chemical View of Density Functional Theory», *J. Phys. Chem. A*, vol. 101 (30), pp. 5383–5403, URL http://pubs.acs.org/doi/abs/10.1021/jp9703768. {Cited on page 120.}

[15] Bartashevich, E., Troitskaya, E., Pendás, Á. M. and Tsirelson, V. (2015), «Understanding the bifurcated halogen bonding NHalN in bidentate diazaheterocyclic compounds», *Comput. Theor. Chem.*, vol. 1053, pp. 229–237, URL http://linkinghub.elsevier.com/retrieve/pii/S2210271X14004307. {Cited on page 121.}

[16] Bartashevich, E. V., Pendás, Á. M. and Tsirelson, V. G. (2014), «An anatomy of intramolecular atomic interactions in halogen-substituted trinitromethanes», *Phys. Chem. Chem. Phys.*, vol. 16 (31), p. 16 780, URL http://xlink.rsc.org/?DOI=C4CP01257G. {Cited on page 121.}

[17] Becke, A. D. (2014), «Perspective: Fifty years of density-functional theory in chemical physics», *J. Chem. Phys.*, vol. 140 (18), p. 18A301, URL http://scitation.aip.org/content/aip/journal/jcp/140/18/10.1063/1.4869598. {Cited on page 44.}

[18] Becke, A. D. and Edgecombe, K. E. (1990), «A simple measure of electron localization in atomic and molecular systems», *J. Chem. Phys.*, vol. 92 (9), pp. 5397–5403, URL http://aip.scitation.org/doi/10.1063/1.458517. {Cited on page 121.}

[19] Becke, A. D. and Johnson, E. R. (2007), «Exchange-hole dipole moment and the dispersion interaction revisited», *J. Chem. Phys.*, vol. 127 (15), p. 154 108, URL http://scitation.aip.org/content/aip/journal/jcp/127/15/10.1063/1.2795701. {Cited on pages 45 and 168.}

[20] Blanco, M. A. (1997), *Métodos cuánticos locales para la simulación de materiales iónicos. Fundamentos, algoritmos y aplicaciones.*, Ph.D. thesis, University of Oviedo. {Cited on page 202.}

[21] Blanco, M. A., Martín Pendás, A. and Francisco, E. (2005), «Interacting Quantum Atoms: A Correlated Energy Decomposition Scheme Based on the Quantum Theory of Atoms in Molecules», *J. Chem. Theory Comput.*, vol. 1 (6), pp. 1096–1109, URL http://pubs.acs.org/doi/abs/10.1021/ct0501093. {Cited on page 147.}

[22] Blöchl, P. E. (1994), «Projector augmented-wave method», *Phys. Rev. B*, vol. 50 (24), pp. 17 953–17 979, URL http://link.aps.org/doi/10.1103/PhysRevB.50.17953. {Cited on pages 61 and 167.}

[23] Blöchl, P. E., Jepsen, O. and Andersen, O. K. (1994), «Improved tetrahedron method for Brillouin-zone integrations», *Phys. Rev. B*, vol. 49 (23), pp. 16 223–16 233, URL http://link.aps.org/doi/10.1103/PhysRevB.49.16223. {Cited on page 60.}

[24] Bonaccorsi, R., Scrocco, E. and Tomasi, J. (1970), «Molecular SCF Calculations for the Ground State of Some Three-Membered Ring Molecules: (CH2)3, (CH2)2NH, (CH2)2NH2+, (CH2)2O, (CH2)2S, (CH)2CH2, and N2CH2», *The Journal of Chemical Physics*, vol. 52 (10), pp. 5270–5284, URL http://scitation.aip.org/content/aip/journal/jcp/52/10/10.1063/1.1672775. {Cited on page 82.}

[25] Borden, W. T. (2017), «Why Are Addition Reactions to N 2 Thermodynamically Unfavorable?», *J. Phys. Chem. A*, vol. 121 (5), pp. 1140–1144, URL http://pubs.acs.org/doi/abs/10.1021/acs.jpca.6b11728. {Cited on page 96.}

[26] Born, M. and Huang, K. (1998), *Dynamical Theory of Crystal Lattices (Oxford Classic Texts in the Physical Sciences)*, Clarendon Press. {Cited on page 9.}

[27] Born, M. and Oppenheimer, R. (1927), «Zur Quantentheorie der Molekeln», *Ann. Phys.*, vol. 389 (20), pp. 457–484, URL http://doi.wiley.com/10.1002/andp.19273892002. {Cited on page 9.}

[28] Bouhmaida, N., Dutheil, M., Ghermani, N. E. and Becker, P. (2002), «Gradient vector field and properties of the experimental electrostatic potential: Application to ibuprofen drug molecule», *J. Chem. Phys.*, vol. 116 (14), pp. 6196–6204. {Cited on page 164.}

[29] Brea, O., El Khatib, M., Angeli, C., Bendazzoli, G. L., Evangelisti, S. and Leininger, T. (2013), «Behavior of the Position–Spread Tensor in Diatomic Systems», *J. Chem. Theory Comput.*, vol. 9 (12), pp. 5286–5295, URL http://pubs.acs.org/doi/abs/10.1021/ct400453b. {Cited on pages 150, 154, and 156.}

[30] Cancès, E., Keriven, R., Lodier, F. and Savin, A. (2004), «How electrons guard the space: shape optimization with probability distribution criteria», *Theor. Chem. Accounts Theory, Comput. Model. (Theoretica Chim. Acta)*, vol. 111 (2-6), pp. 373–380, URL http://link.springer.com/10.1007/s00214-003-0509-4. {Cited on page 69.}

[31] Cao, W., Gatti, C., MacDougall, P. and Bader, R. (1987), «On the presence of non-nuclear attractors in the charge distributions of Li and Na clusters», *Chem. Phys. Lett.*, vol. 141 (5), pp. 380–385, URL http://linkinghub.elsevier.com/retrieve/pii/0009261487850443. {Cited on page 147.}

[32] Causà, M. and Savin, A. (2011), «Maximum Probability Domains in Crystals: The Rock-Salt Structure», *J. Phys. Chem. A*, vol. 115 (45), pp. 13 139–13 148, URL http://pubs.acs.org/doi/abs/10.1021/jp205622x. {Cited on page 120.}

[33] Chauvin, R., Lepetit, C., Silvi, B. and Alikhani, E., eds. (2016), *Applications of Topological Methods in Molecular Chemistry*, Springer International Publishing, Cham, URL http://link.springer.com/10.1007/978-3-319-29022-5. {Cited on page 121.}

[34] Chisholm, C. D. H. (1976), *Group theoretical techniques in quantum chemistry*, Academic Press. {Cited on page 204.}

[35] Clark, A. E. and Davidson, E. R. (2001), «Local spin», *J. Chem. Phys.*, vol. 115 (16), pp. 7382–7392, URL http://aip.scitation.org/doi/10.1063/1.1407276. {Cited on page 100.}

[36] Coleman, A. J. (1963), «Structure of Fermion Density Matrices», *Rev. Mod. Phys.*, vol. 35 (3), pp. 668–686, URL http://link.aps.org/doi/10.1103/RevModPhys.35.668. {Cited on page 20.}

[37] Condon, E. U. (1930), «The Theory of Complex Spectra», *Phys. Rev.*, vol. 36 (7), pp. 1121–1133, URL http://link.aps.org/doi/10.1103/PhysRev.36.1121. {Cited on page 24.}

[38] Coppens, P. (1997), *X-Ray Charge Densities and Chemical Bonding*, International Union of Crystallography Texts on Crystallography, Oxford University Press, USA. {Cited on pages 53 and 54.}

[39] Cremer, D., Wu, A., Larsson, A. and Kraka, E. (2000), «Some Thoughts about Bond Energies, Bond Lengths, and Force Constants», *J. Mol.*

Model., vol. 6 (4), pp. 396–412, URL http://link.springer.com/10.1007/PL00010739. {Cited on page 96.}

[40] Cuevas, J. C. and Scheer, E. (2010), *Molecular Electronics. An Introduction to Theory and Experiment*, WORLD SCIENTIFIC, URL http://www.worldscientific.com/worldscibooks/10.1142/7434. {Cited on page 146.}

[41] Danovich, D., Shaik, S., Rzepa, H. S. and Hoffmann, R. (2013), «A Response to the Critical Comments on "One Molecule, Two Atoms, Three Views, Four Bonds?"», *Angew. Chemie Int. Ed.*, vol. 52 (23), pp. 5926–5928, URL http://doi.wiley.com/10.1002/anie.201302350. {Cited on page 96.}

[42] Delle Site, L. (2004), «A Note on the Initial Condition of the Differential Equation Which Defines Proper Quantum Topological Subspaces», *J. Math. Chem.*, vol. 35 (3), pp. 289–295, URL http://link.springer.com/10.1023/B:JOMC.0000033260.32364.bb. {Cited on page 78.}

[43] Desiraju, G. R., Vittal, J. J. and Ramanan, A. (2011), *Crystal Engineering. A textbook*, Wold Scientific. Co-Published with Indian Institute of Science (IISc), Bangalore, India, URL http://www.worldscientific.com/worldscibooks/10.1142/8060. {Cited on page 146.}

[44] El Khatib, M., Leininger, T., Bendazzoli, G. L. and Evangelisti, S. (2014), «Computing the Position-Spread tensor in the CAS-SCF formalism», *Chem. Phys. Lett.*, vol. 591, pp. 58–63, URL http://dx.doi.org/10.1016/j.cplett.2013.10.080http://linkinghub.elsevier.com/retrieve/pii/S0009261413013602. {Cited on page 150.}

[45] El Khatib, M., Brea, O., Fertitta, E., Bendazzoli, G. L., Evangelisti, S. and Leininger, T. (2015), «The total position-spread tensor: Spin partition», *J. Chem. Phys.*, vol. 142 (9), p. 094 113, URL http://aip.scitation.org/doi/10.1063/1.4913734. {Cited on page 150.}

[46] Espinosa, E., Molins, E. and Lecomte, C. (1997), «Electron density study of the one-dimensional organic metal bis(thiodimethylene)-tetrathiafulvalene tetracyanoquinodimethane», *Phys. Rev. B*, vol. 56, pp. 1820–1833, URL http://link.aps.org/doi/10.1103/PhysRevB.56.1820. {Cited on pages 166 and 185.}

[47] Espinosa, E., Molins, E. and Lecomte, C. (1997), «Electron density study of the one-dimensional organic metal bis(thiodimethylene)-tetrathiafulvalene tetracyanoquinodimethane», *Phys. Rev. B*, vol. 56 (4), pp. 1820–1833, URL http://link.aps.org/doi/10.1103/PhysRevB.56.1820. {Cited on pages 167 and 182.}

[48] Foroutan-Nejad, C., Badri, Z. and Marek, R. (2015), «Multi-center covalency: revisiting the nature of anion–π interactions», *Phys. Chem. Chem. Phys.*, vol. 17 (45), pp. 30 670–30 679, URL http://xlink.rsc.org/?DOI=C5CP05777A. {Cited on page 141.}

[49] Francisco, E. and Costales, A. (2015), «An energy partition method based on localized molecular orbitals», *Comput. Theor. Chem.*, vol. 1053, pp. 77–84, URL http://linkinghub.elsevier.com/retrieve/pii/S2210271X14004083. {Cited on page 123.}

[50] Francisco, E. and Martín Pendás, A. (2014), «Electron number distribution functions from molecular wavefunctions. Version 2», *Comput. Phys. Commun.*, vol. 185 (10), pp. 2663–2682, URL http://dx.doi.org/10.1016/j.cpc.2014.05.009http://linkinghub.elsevier.com/retrieve/pii/S0010465514001635. {Cited on page 104.}

[51] Francisco, E., Martín Pendás, A. and Blanco, M. A. (2006), «A Molecular Energy Decomposition Scheme for Atoms in Molecules», *J. Chem. Theory Comput.*, vol. 2 (1), pp. 90–102, URL http://pubs.acs.org/doi/abs/10.1021/ct0502209. {Cited on pages 120, 121, and 122.}

[52] Francisco, E., Martín Pendás, A. and Blanco, M. A. (2007), «Electron number probability distributions for correlated wave functions», *J. Chem. Phys.*, vol. 126 (9), p. 094 102, URL http://scitation.aip.org/content/aip/journal/jcp/126/9/10.1063/1.2709883. {Cited on page 156.}

[53] Francisco, E., Martín Pendás, A. and Blanco, M. A. (2009), «A connection between domain-averaged Fermi hole orbitals and electron number distribution functions in real space», *J. Chem. Phys.*, vol. 131 (12), p. 124 125, URL http://aip.scitation.org/doi/10.1063/1.3239467. {Cited on page 152.}

[54] Frenking, G. and Hermann, M. (2013), «Critical Comments on "One Molecule, Two Atoms, Three Views, Four Bonds?"», *Angew. Chemie Int. Ed.*, vol. 52 (23), pp. 5922–5925, URL http://doi.wiley.com/10.1002/anie.201301485. {Cited on pages 96 and 104.}

[55] Frenking, G. and Tonner, R. (2009), «Divalent carbon(0) compounds», *Pure Appl. Chem.*, vol. 81 (4), pp. 597–614, URL http://www.degruyter.com/view/j/pac.2009.81.issue-4/pac-con-08-11-03/pac-con-08-11-03.xml. {Cited on page 96.}

[56] Frenking, G., Tonner, R., Klein, S., Takagi, N., Shimizu, T., Krapp, A., Pandey, K. K. and Parameswaran, P. (2014), «New bonding modes of carbon and heavier group 14 atoms Si–Pb», *Chem. Soc. Rev.*, vol. 43 (14), pp. 5106–5139, URL http://www.ncbi.nlm.nih.gov/pubmed/24916774http://xlink.rsc.org/?DOI=C4CS00073K. {Cited on page 96.}

[57] Fulton, T. A. and Dolan, G. J. (1987), «Observation of single-electron charging effects in small tunnel junctions», *Phys. Rev. Lett.*, vol. 59 (1), pp. 109–112, URL http://link.aps.org/doi/10.1103/PhysRevLett.59.109. {Cited on page 146.}

[58] Gadre, S. R. and Bendale, R. D. (1986), «On the similarity between molecular electron densities, electrostatic potentials and bare nuclear potentials», *Chemical Physics Letters*, vol. 130 (6), pp. 515 – 521, URL http://www.sciencedirect.com/science/article/pii/0009261486802494. {Cited on page 83.}

[59] Gadre, S. R. and Bendale, R. D. (1986), «On the similarity between molecular electron densities, electrostatic potentials and bare nuclear potentials», *Chem. Phys. Lett.*, vol. 130 (6), pp. 515–521, URL http://linkinghub.elsevier.com/retrieve/pii/0009261486802494. {Cited on page 164.}

[60] Gadre, S. R., Kulkarni, S. a. and Shrivastava, I. H. (1992), «Molecular electrostatic potentials: A topographical study», *J. Chem. Phys.*, vol. 96 (7), p. 5253, URL http://scitation.aip.org/content/aip/journal/jcp/96/7/10.1063/1.462710. {Cited on page 164.}

[61] Gallo-Bueno, A., Francisco, E. and Martin Pendas, A. (2016), «Decay rate of real space delocalization measures: a comparison between analytical and test systems», *Phys. Chem. Chem. Phys.*, vol. 18 (17), pp. 11 772–11 780, URL http://pubs.rsc.org/en/content/articlehtml/2016/cp/c5cp06098bhttp://xlink.rsc.org/?DOI=C5CP06098B. {Cited on pages 149, 153, and 157.}

[62] Gallo-Bueno, A., Kohout, M. and Martin Pendás, A. (2016), «Decay Rate of Correlated Real-Space Delocalization Measures: Insights into Chemical Bonding and Mott Transitions from Hydrogen Chains», *J. Chem. Theory Comput.*, vol. 12 (7), pp. 3053–3062, URL http://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b00139. {Cited on pages 149, 150, 153, 157, and 158.}

[63] García-Revilla, M., Francisco, E., Popelier, P. L. A. and Martín Pendás, A. (2013), «Domain-Averaged Exchange-Correlation Energies as a Physical Underpinning for Chemical Graphs», *ChemPhysChem*, vol. 14 (6), pp. 1211–1218, URL http://doi.wiley.com/10.1002/cphc.201300092. {Cited on pages 121 and 123.}

[64] Garcia-Revilla, M., Popelier, P. L. A., Francisco, E. and Martin Pendas, A. (2011), «Nature of Chemical Interactions from the Profiles of Electron Delocalization Indices», *J. Chem. Theory Comput.*, vol. 7 (6), pp. 1704–1711, URL http://pubs.acs.org/doi/abs/10.1021/ct2001842. {Cited on page 121.}

[65] Garcia-Revilla, M., Popelier, P. L. A., Francisco, E. and Martin Pendas, A. (2011), «Nature of Chemical Interactions from the Profiles of Electron Delocalization Indices», *J. Chem. Theory Comput.*, vol. 7 (6), pp. 1704–1711, URL http://pubs.acs.org/doi/abs/10.1021/ct2001842. {Cited on pages 149 and 157.}

[66] Gatti, C. and Macchi, P., eds. (2012), *Modern Charge-Density Analysis*, Springer. {Cited on page 1.}

[67] Gatti, C. and Macchi, P., eds. (2012), *Modern Charge-Density Analysis*, Springer Netherlands, Dordrecht, URL http://link.springer.com/10.1007/978-90-481-3836-4. {Cited on pages 112 and 121.}

[68] Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G. L., Cococcioni, M., Dabo, I., Dal Corso, A., de Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A. P., Smogunov, A., Umari, P. and Wentzcovitch, R. M. (2009), «QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials.», *J. Phys. Condens. Matter*, vol. 21 (39), p. 395 502. {Cited on page 168.}

[69] Gimarc, B. (1979), *Molecular structure and bonding: the qualitative molecular orbital approach*, Academic Press. {Cited on pages 120 and 147.}

[70] Grunenberg, J. (2012), «Quantum chemistry: Quadruply bonded carbon», *Nat. Chem.*, vol. 4 (3), pp. 154–155, URL http://dx.doi.org/10.1038/nchem.1274http://www.nature.com/doifinder/10.1038/nchem.1274. {Cited on page 96.}

[71] Grunenberg, J. (2017), «Ill-defined chemical concepts: The problem of quantification», *Int. J. Quantum Chem.*, (October 2016), p. e25 359, URL http://doi.wiley.com/10.1002/qua.25359. {Cited on page 97.}

[72] G.Tsirelson, V., Avilov, a. S., Lepeshov, G. G., Kulygin, a. K., Stahn, J., Pietsch, U. and Spence, J. C. H. (2001), «Quantitative Analysis of the Electrostatic Potential in Rock-Salt Crystals Using Accurate Electron Diffraction Data», *J. Phys. Chem.*, vol. 105, pp. 5068–5074. {Cited on page 164.}

[73] Guevara-Vela, J. M., Chávez-Calvillo, R., García-Revilla, M., Hernández-Trujillo, J., Christiansen, O., Francisco, E., Martín Pendás, Á. and Rocha-Rinza, T. (2013), «Hydrogen-Bond Cooperative Effects in Small Cyclic Water Clusters as Revealed by the Interacting Quantum Atoms Approach», *Chem. - A Eur. J.*, vol. 19 (42), pp. 14 304–14 315, URL http://doi.wiley.com/10.1002/chem.201300656. {Cited on page 121.}

[74] Guillot, B., Viry, L., Guillot, R., Lecomte, C. and Jelsch, C. (2001), «Refinement of proteins at subatomic resolution with MOPRO», *J. Appl. Crystallogr.*, vol. 34 (2), pp. 214–223. {Cited on page 167.}

[75] Hahn, T., ed. (2002), *International tables for crystallography, vol. A*, International Union of Crystallography by Kluwer Academic Publishers. {Cited on page 48.}

[76] Hansen, N. K. and Coppens, P. (1978), «Testing aspherical atom refinements on small-molecule data sets», *Acta Crystallographica Section A*, vol. 34 (6), pp. 909–921, URL http://dx.doi.org/10.1107/S0567739478001886. {Cited on page 53.}

[77] Hansen, N. K. and Coppens, P. (1978), «Testing aspherical atom refinements on small-molecule data sets», *Acta Crystallogr. Sect. A Cryst. Physics, Diffraction, Theor. Gen. Crystallogr.*, vol. 34 (6), pp. 909–921, URL http://scripts.iucr.org/cgi-bin/paper?S0567739478001886. {Cited on page 167.}

[78] Henkelman, G., Arnaldsson, A. and Jónsson, H. (2006), «A fast and robust algorithm for Bader decomposition of charge density», *Comput. Mater. Sci.*, vol. 36 (3), pp. 354–360, URL http://linkinghub.elsevier.com/retrieve/pii/S0927025605001849. {Cited on page 168.}

[79] Hermann, M. and Frenking, G. (2016), «The Chemical Bond in C 2», *Chem. - A Eur. J.*, vol. 22 (12), pp. 4100–4108, URL http://doi.wiley.com/10.1002/chem.201503762.

[80] Herring, C. (1940), «A New Method for Calculating Wave Functions in Crystals», *Phys. Rev.*, vol. 57 (12), pp. 1169–1177, URL http://link.aps.org/doi/10.1103/PhysRev.57.1169. {Cited on page 58.}

[81] Herrmann, C., Reiher, M. and Hess, B. A. (2005), «Comparative analysis of local spin definitions», *J. Chem. Phys.*, vol. 122 (3), p. 034 102, URL http://aip.scitation.org/doi/10.1063/1.1829050. {Cited on page 100.}

[82] Hohenberg, P. and Kohn, W. (1964), «Inhomogeneous Electron Gas», *Phys. Rev.*, vol. 136 (3B), pp. B864–B871, URL http://link.aps.org/doi/10.1103/PhysRev.136.B864. {Cited on page 39.}

[83] Hylleraas, E. A. (1929), «Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium», *Zeitschrift für Phys.*, vol. 54 (5-6), pp. 347–366, URL http://link.springer.com/10.1007/BF01375457. {Cited on page 13.}

[84] Hylleraas, E. A. and Undheim, B. (1930), «Numerische Berechnung der 2S-Terme von Ortho- und Par-Helium», *Zeitschrift fur Phys.*, vol. 65 (11-12), pp. 759–772, URL http://link.springer.com/10.1007/BF01397263. {Cited on page 26.}

[85] Jelsch, C., Guillot, B., Lagoutte, A. and Lecomte, C. (2005), «Advances in protein and small-molecule charge-density refinement methods using *MoPro*», *Journal of Applied Crystallography*, vol. 38 (1), pp. 38–54, URL http://dx.doi.org/10.1107/S0021889804025518.

[86] Jelsch, C., Guillot, B., Lagoutte, A. and Lecomte, C. (2005), «Advances in protein and small-molecule charge-density refinement methods using MoPro», *J. Appl. Crystallogr.*, vol. 38 (1), pp. 38–54. {Cited on page 167.}

[87] Kato, T. (1957), «On the eigenfunctions of many-particle systems in quantum mechanics», *Commun. Pure Appl. Math.*, vol. 10 (2), pp. 151–177, URL http://doi.wiley.com/10.1002/cpa.3160100201. {Cited on page 13.}

[88] Kohn, W. (1964), «Theory of the Insulating State», *Phys. Rev.*, vol. 133 (1A), pp. A171–A181, URL http://link.aps.org/doi/10.1103/PhysRev.133.A171. {Cited on pages 147, 148, and 196.}

[89] Kohn, W. and Sham, L. J. (1965), «Self-Consistent Equations Including Exchange and Correlation Effects», *Phys. Rev.*, vol. 140 (4A), pp. A1133–A1138, URL http://link.aps.org/doi/10.1103/PhysRev.140.A1133. {Cited on page 120.}

[90] Kohout, M. (2004), «A measure of electron localizability», *Int. J. Quantum Chem.*, vol. 97 (1), pp. 651–658, URL http://doi.wiley.com/10.1002/qua.10768. {Cited on pages 120 and 147.}

[91] Koopmans, T. (1934), «Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms», *Physica*, vol. 1 (1-6), pp. 104–113, URL http://linkinghub.elsevier.com/retrieve/pii/S0031891434900112. {Cited on page 26.}

[92] Kosov, D. S. and Popelier, P. L. A. (2000), «Convergence of the multipole expansion for electrostatic potentials of finite topological atoms», *J. Chem. Phys.*, vol. 113 (10), pp. 3969–3974, URL http://aip.scitation.org/doi/10.1063/1.1288384. {Cited on page 125.}

[93] Kresse, G. and Furthmüller, J. (1996), «Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set», *Phys. Rev. B*, vol. 54 (16), pp. 11 169–11 186, URL http://link.aps.org/doi/10.1103/PhysRevB.54.11169. {Cited on page 167.}

[94] Krukau, A. V., Vydrov, O. A., Izmaylov, A. F. and Scuseria, G. E. (2006), «Influence of the exchange screening parameter on the performance of screened hybrid functionals», *J. Chem. Phys.*, vol. 125 (22), p. 224 106, URL http://scitation.aip.org/content/aip/journal/jcp/125/22/10.1063/1.2404663. {Cited on page 167.}

[95] Kudinov, E. K. (1991), «Fisica Tverdogo Tela 33, 2306 (1991)[English translation]», *Sov. Phys. Solid State*, vol. 33, p. 1299, URL http://link.springer.com/10.1134/1.1131030. {Cited on page 148.}

[96] Kurth, S. and Perdew, J. P. (2000), «Role of the exchange-correlation energy: Nature's glue», *Int. J. Quantum Chem.*, vol. 77 (5), pp. 814–818, URL http://doi.wiley.com/10.1002/{%}28SICI{%}291097-461X{%}282000{%}2977{%}3A5{%}3C814{%}3A{%}3AAID-QUA3{%}3E3.0.CO{%}3B2-F. {Cited on page 120.}

[97] Leboeuf, M., Köster, A. M., Jug, K. and Salahub, D. R. (1999), «Topological analysis of the molecular electrostatic potential», *J. Chem. Phys.*, vol. 111 (11), pp. 4893–4905, URL http://jcp.aip.org/resource/1/jcpsa6/v111/i11/p4893{_}s1?isAuthorized=no. {Cited on page 164.}

[98] Lewin, M. (2004), «Solutions of the Multiconfiguration Equations in Quantum Chemistry», *Arch. Ration. Mech. Anal.*, vol. 171 (1), pp. 83–114, URL http://link.springer.com/10.1007/s00205-003-0281-6. {Cited on page 37.}

[99] Li, L. and Parr, R. G. (1986), «The atom in a molecule: A density matrix approach», *J. Chem. Phys.*, vol. 84 (3), p. 1704, URL http://scitation.aip.org/content/aip/journal/jcp/84/3/10.1063/1.450468. {Cited on page 87.}

[100] Long, L. H. (1951), «The valence-state energy of the tetravalent carbon atom», *Experientia*, vol. 7 (5), pp. 195–200, URL http://link.springer.com/10.1007/BF02148916. {Cited on pages 96 and 104.}

[101] Long, L. H. and Sackman, J. F. (1958), «The heat of formation and physical properties of gallium trimethyl», *Trans. Faraday Soc.*, vol. 54, p. 1797, URL http://pubs.rsc.org/en/content/articlehtml/1958/tf/tf9585401797http://xlink.rsc.org/?DOI=tf9585401797. {Cited on page 96.}

[102] Löwdin, P.-O. (1955), «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 (6), pp. 1474–1489, URL http://link.aps.org/doi/10.1103/PhysRev.97.1474. {Cited on pages 13, 16, and 17.}

[103] MacDonald, J. K. L. (1933), «Successive Approximations by the Rayleigh-Ritz Variation Method», *Phys. Rev.*, vol. 43 (10), pp. 830–833, URL http://link.aps.org/doi/10.1103/PhysRev.43.830. {Cited on page 26.}

[104] Makov, G. and Payne, M. C. (1995), «Periodic boundary conditions in ab initio calculations», *Phys. Rev. B*, vol. 51 (7), pp. 4014–4022, URL http://link.aps.org/doi/10.1103/PhysRevB.51.4014. {Cited on page 54.}

[105] Malcolm, N. O. J. and Popelier, P. L. A. (2003), «The full topology of the Laplacian of the electron density: scrutinising a physical basis for the VSEPR model», *Faraday Discuss.*, vol. 124, pp. 353–363, URL http://dx.doi.org/10.1039/B211650M. {Cited on page 82.}

[106] Martin Pendas, A., Blanco, M. and Francisco, E. (2006), «Revisiting the variational nature of the quantum theory of atoms in molecules», *Chem. Phys. Lett.*, vol. 417 (1-3), pp. 16–21, URL http://linkinghub.elsevier.com/retrieve/pii/S0009261406007512http://linkinghub.elsevier.com/retrieve/pii/S0009261405015058. {Cited on pages 78 and 87.}

[107] Martín Pendás, A. and Francisco, E. (????), «Promolden: A QTAIM/IQA code (Available from the authors upon request by writing to ampendas@uniovi.es)», . {Cited on pages 128 and 154.}

[108] Martín Pendás, A., Francisco, E. and Blanco, M. (2007), «Spatial localization, correlation, and statistical dependence of electrons in atomic domains: The X1Sg+ and b3Su+ states of H2», *Chem. Phys. Lett.*, vol. 437 (4-6), pp. 287–292, URL http://linkinghub.elsevier.com/retrieve/pii/S0009261407002126. {Cited on page 154.}

[109] Martín Pendás, A., Francisco, E. and Blanco, M. A. (2006), «Binding Energies of First Row Diatomics in the Light of the Interacting Quantum Atoms Approach», *J. Phys. Chem. A*, vol. 110 (47), pp. 12 864–12 869, URL http://pubs.acs.org/doi/abs/10.1021/jp063607w. {Cited on pages 121, 123, and 149.}

[110] Martín Pendás, A., Francisco, E. and Blanco, M. a. (2007), «An electron number distribution view of chemical bonds in real space», *Phys. Chem. Chem. Phys.*, vol. 9 (9), pp. 1087–1092, URL http://xlink.rsc.org/?DOI=B616310F. {Cited on pages 101 and 102.}

[111] Martín Pendás, A., Francisco, E. and Blanco, M. a. (2007), «Charge transfer, chemical potentials, and the nature of functional groups: answers from quantum chemical topology», *Faraday Discuss.*, vol. 135, pp. 423–438, URL http://xlink.rsc.org/?DOI=B604983D. {Cited on page 156.}

[112] Martin Pendas, A., Francisco, E. and Blanco, M. A. (2007), «Spin resolved electron number distribution functions: How spins couple in real space», *J. Chem. Phys.*, vol. 127 (14), p. 144 103, URL http://scitation.aip.org/content/aip/journal/jcp/127/14/10.1063/1.2784392. {Cited on pages 71, 100, and 102.}

[113] Mata, I., Molins, E. and Espinosa, E. (2007), «Zero-Flux Surfaces of the Electrostatic Potential: The Border of Influence Zones of Nucleophilic and Electrophilic Sites in Crystalline Environment», *The Journal of Physical Chemistry A*, vol. 111 (39), pp. 9859–9870, URL http://pubs.acs.org/doi/abs/10.1021/jp074032l, pMID: 17727276. {Cited on pages 2, 83, and 86.}

[114] Mata, I., Molins, E. and Espinosa, E. (2007), «Zero-flux surfaces of the electrostatic potential: The border of influence zones of nucleophilic and electrophilic sites in crystalline environment», *J. Phys. Chem. A*, vol. 111 (39), pp. 9859–9870. {Cited on page 164.}

[115] Mata, I., Molins, E., Alkorta, I. and Espinosa, E. (2007), «Topological properties of the electrostatic potential in weak and moderate N\cdotsH hydrogen bonds», *J. Phys. Chem. A*, vol. 111 (28), pp. 6425–6433. {Cited on page 84.}

[116] Mata, I., Molins, E., Alkorta, I. and Espinosa, E. (2007), «Topological Properties of the Electrostatic Potential in Weak and Moderate N···H Hydrogen Bonds», *The Journal of Physical Chemistry A*, vol. 111 (28), pp. 6425–6433, URL http://pubs.acs.org/doi/abs/10.1021/jp071924c, pMID: 17583329. {Cited on page 2.}

[117] Matta, C. and Boyd, R., eds. (2007), *The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design*, Wiley. {Cited on page 1.}

[118] Mayer, I. (1983), «Charge, bond order and valence in the AB initio SCF theory», *Chem. Phys. Lett.*, vol. 97 (3), pp. 270–274, URL http://linkinghub.elsevier.com/retrieve/pii/0009261483800050. {Cited on page 121.}

[119] Mayer, I. (2007), «Bond order and valence indices: A personal account», *J. Comput. Chem.*, vol. 28 (1), pp. 204–221, URL http://doi.wiley.com/10.1002/jcc.20494. {Cited on page 149.}

[120] Mayer, I. (2007), «Local spins: An alternative treatment for single determinant wave functions», *Chem. Phys. Lett.*, vol. 440 (4-6), pp. 357–359, URL http://linkinghub.elsevier.com/retrieve/pii/S0009261407005064. {Cited on page 100.}

[121] Mayer, I. (2009), «Local spins: An improved treatment for correlated wave functions», *Chem. Phys. Lett.*, vol. 478 (4-6), pp. 323–326, URL http://dx.doi.org/10.1016/j.cplett.2009.07.087http://linkinghub.elsevier.com/retrieve/pii/S0009261409009191. {Cited on page 100.}

[122] Mayer, I. and Hamza, A. (2003), «Interatomic exchange energy components», *Int. J. Quantum Chem.*, vol. 92 (2), pp. 174–180, URL http://doi.wiley.com/10.1002/qua.10504. {Cited on page 123.}

[123] Mazziotti, D. A. (2011), «Large-scale semidefinite programming for many-electron quantum mechanics», *Phys. Rev. Lett.*, vol. 106 (8), pp. 7–10. {Cited on page 20.}

[124] Mazziotti, D. A. (2012), «Structure of Fermionic density matrices: Complete N-representability conditions», *Phys. Rev. Lett.*, vol. 108 (26), pp. 1–5. {Cited on page 20.}

[125] Mazziotti, D. A. (2012), «Two-Electron Reduced Density Matrix as the Basic Variable in Many-Electron Quantum Chemistry and Physics», *Chem. Rev.*, vol. 112 (1), pp. 244–262, URL http://pubs.acs.org/doi/abs/10.1021/cr2000493. {Cited on page 20.}

[126] McWeeny, R. (1960), «Some Recent Advances in Density Matrix Theory», *Rev. Mod. Phys.*, vol. 32 (2), pp. 335–369, URL http://link.aps.org/doi/10.1103/RevModPhys.32.335. {Cited on page 16.}

[127] McWeeny, R. (1997), «Separability in Quantum Mechanics», *Mol. Eng.*, vol. 7 (1/2), pp. 7–26, URL http://dx.doi.org/10.1023/A:1008266011351http://link.springer.com/10.1023/A:1008266011351. {Cited on page 87.}

[128] Methfessel, M. and Paxton, A. T. (1989), «High-precision sampling for Brillouin-zone integration in metals», *Phys. Rev. B*, vol. 40 (6), pp. 3616–3621, URL http://link.aps.org/doi/10.1103/PhysRevB.40.3616. {Cited on page 60.}

[129] Murray, J. and Sen, K. (1996), *Molecular Electrostatic Potentials: Concepts and Applications*, Theoretical and Computational Chemistry, Elsevier Science. {Cited on page 83.}

[130] Nasertayoob, P. and Shahbazian, S. (2008), «Revisiting the foundations of quantum theory of atoms in molecules (QTAIM): The variational procedure and the zero-flux conditions», *Int. J. Quantum Chem.*, vol. 108 (9), pp. 1477–1484, URL http://doi.wiley.com/10.1002/qua.21665. {Cited on page 78.}

[131] Netzel, J. and van Smaalen, S. (2009), «Topological properties of hydrogen bonds and covalent bonds from charge densities obtained by the maximum entropy method (MEM)», *Acta Crystallographica Section B*, vol. 65 (5), pp. 624–638, URL http://dx.doi.org/10.1107/S0108768109026767. {Cited on page 51.}

[132] Novák, M., Foroutan-Nejad, C. and Marek, R. (2016), «Modulating Electron Sharing in Ion-π-Receptors via Substitution and External Electric Field: A Route toward Bond Strengthening», *J. Chem. Theory Comput.*, vol. 12 (8), pp. 3788–3795, URL http://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b00586. {Cited on page 141.}

[133] Omar, Y. (2005), «Indistinguishable particles in quantum mechanics: an introduction», *Contemp. Phys.*, vol. 46 (6), pp. 437–448, URL http://www.tandfonline.com/doi/abs/10.1080/00107510500361274. {Cited on page 11.}

[134] Otero-de-la Roza, A., Johnson, E. R. and Luaña, V. (2014), «Critic2: A program for real-space analysis of quantum chemical interactions in solids», *Computer Physics Communications*, vol. 185 (3), pp. 1007 – 1018, URL http://www.sciencedirect.com/science/article/pii/S0010465513003718.

[135] Otero-de-la Roza, A., Johnson, E. R. and Luaña, V. (2014), «Critic2: A program for real-space analysis of quantum chemical interactions in solids», *Comput. Phys. Commun.*, vol. 185 (3), pp. 1007–1018, URL http://dx.doi.org/10.1016/j.cpc.2013.10.026http://linkinghub.elsevier.com/retrieve/pii/S0010465513003718. {Cited on page 167.}

[136] Otero-de-la Roza, A. and Luaña, V. (2010), «Topological Characterization of the Electron Density Laplacian in Crystals. The Case of the Group IV Elements», *Journal of Chemical Theory and Computation*, vol. 6 (12), pp. 3761–3779, URL http://pubs.acs.org/doi/abs/10.1021/ct100269e. {Cited on page 82.}

[137] Otero-de-la Roza, A., Blanco, M. A., Martín Pendás, A. and Luaña, V. (2009), «Critic: a new program for the topological analysis of solid-state electron densities», *Comput. Phys. Commun.*, vol. 180, pp. 157–166.

[138] Otero-de-la Roza, A., Blanco, M., Pendás, A. M. and Luaña, V. (2009), «Critic: a new program for the topological analysis of solid-state electron densities», *Comput. Phys. Commun.*, vol. 180 (1), pp. 157–166, URL http://linkinghub.elsevier.com/retrieve/pii/S0010465508002865. {Cited on page 167.}

[139] Otero-de-la Roza, A., Cao, B. H., Price, I. K., Hein, J. E. and Johnson, E. R. (2014), «Predicting the Relative Solubilities of Racemic and Enantiopure Crystals by Density-Functional Theory», *Angew. Chemie Int. Ed.*, vol. 53 (30), pp. 7879–7882, URL http://doi.wiley.com/10.1002/anie.201403541. {Cited on page 45.}

[140] Pack, J. D. and Monkhorst, H. J. (1977), «"Special points for Brillouin-zone integrations"—a reply», *Phys. Rev. B*, vol. 16 (4), pp. 1748–1749, URL http://link.aps.org/doi/10.1103/PhysRevB.16.1748. {Cited on page 60.}

[141] Pathak, R. K. and Gadre, S. R. (1990), «Maximal and minimal characteristics of molecular electrostatic potentials», *The Journal of Chemical Physics*, vol. 93 (3), pp. 1770–1773, URL http://scitation.aip.org/content/aip/journal/jcp/93/3/10.1063/1.459703. {Cited on page 83.}

[142] Pendás, A. M., Blanco, M. A. and Francisco, E. (2004), «Two-electron integrations in the quantum theory of atoms in molecules», *J. Chem. Phys.*, vol. 120 (10), pp. 4581–4592, URL http://aip.scitation.org/doi/10.1063/1.1645788. {Cited on pages 124 and 128.}

[143] Pendás, A. M., Blanco, M. A. and Francisco, E. (2007), «Chemical fragments in real space: Definitions, properties, and energetic decompositions», *J. Comput. Chem.*, vol. 28 (1), pp. 161–184, URL http://doi.wiley.com/10.1002/jcc.20469. {Cited on pages 120 and 123.}

[144] Pendás, A. M., Blanco, M. A. and Francisco, E. (2009), «Steric repulsions, rotation barriers, and stereoelectronic effects: A real space perspective», *J. Comput. Chem.*, vol. 30 (1), pp. 98–109, URL http://doi.wiley.com/10.1002/jcc.21034. {Cited on pages 121 and 149.}

[145] Pendás, A. M., Francisco, E. and Blanco, M. A. (2005), «Two-electron integrations in the Quantum Theory of Atoms in Molecules with correlated wave functions», *J. Comput. Chem.*, vol. 26 (4), pp. 344–351, URL http://doi.wiley.com/10.1002/jcc.20173. {Cited on page 104.}

[146] Pendás, A. M., Francisco, E. and Blanco, M. A. (2005), «Two-electron integrations in the Quantum Theory of Atoms in Molecules with correlated wave functions», *J. Comput. Chem.*, vol. 26 (4), pp. 344–351, URL http://doi.wiley.com/10.1002/jcc.20173. {Cited on pages 122, 124, 128, and 200.}

[147] Pendás, A. M., Francisco, E. and Blanco, M. A. (2007), «Pauling Resonant Structures in Real Space through Electron Number Probability Distributions», *J. Phys. Chem. A*, vol. 111 (6), pp. 1084–1090, URL http://pubs.acs.org/doi/abs/10.1021/jp064600h. {Cited on page 101.}

[148] Pendás, A. M., Francisco, E., Blanco, M. A. and Gatti, C. (2007), «Bond Paths as Privileged Exchange Channels», *Chem. - A Eur. J.*, vol. 13 (33), pp. 9362–9371, URL http://doi.wiley.com/10.1002/chem.200700408. {Cited on pages 121 and 123.}

[149] Perdew, J. P., Burke, K. and Ernzerhof, M. (1996), «Generalized Gradient Approximation Made Simple», *Phys. Rev. Lett.*, vol. 77 (18), pp. 3865–3868, URL http://www.ncbi.nlm.nih.gov/pubmed/10062328{%}5Cnhttp://link.aps.org/doi/10.1103/PhysRevLett.77.3865{%}5Cnhttp://link.aps.org/abstract/PRL/v77/p3865{%}5Cnhttp://link.aps.org/doi/10.1103/PhysRevLett.77.3865http://link.aps.org/doi/10.1103/PhysRevLett.77.3865. {Cited on page 168.}

[150] Phillips, J. C. and Kleinman, L. (1959), «New Method for Calculating Wave Functions in Crystals and Molecules», *Phys. Rev.*, vol. 116 (2), pp. 287–294, URL http://link.aps.org/doi/10.1103/PhysRev.116.287. {Cited on page 59.}

[151] Politzer, P. and Murray, J. S. (2002), «The fundamental nature and role of the electrostatic potential in atoms and molecules», *Theor. Chem. Acc.*, vol. 108 (3), pp. 134–142. {Cited on page 164.}

[152] Ponec, R. (1997), «Electron pairing and chemical bonds. Chemical structure, valences and structural similarities from the analysis of the Fermi holes», *J. Math. Chem.*, vol. 21 (3), pp. 323–333, URL http://link.springer.com/10.1023/A:1019186806180. {Cited on pages 152 and 197.}

[153] Ponec, R. (1998), «Electron pairing and chemical bonds. Molecular structure from the analysis of pair densities and related quantities», *J. Math. Chem.*, vol. 23 (1/2), pp. 85–103, URL http://link.springer.com/10.1023/A:1019160922535. {Cited on page 197.}

[154] Popelier, P. L. A. and Brémond, É. A. G. (2009), «Geometrically faithful homeomorphisms between the electron density and the bare nuclear potential», *Int. J. Quantum Chem.*, vol. 109 (11), pp. 2542–2553, URL http://doi.wiley.com/10.1002/qua.22215. {Cited on pages 121 and 147.}

[155] Popelier, P. L. A., Joubert, L. and Kosov, D. S. (2001), «Convergence of the Electrostatic Interaction Based on Topological Atoms», *J. Phys. Chem. A*, vol. 105 (35), pp. 8254–8261, URL http://pubs.acs.org/doi/abs/10.1021/jp011511q. {Cited on page 125.}

[156] Popelier, P. L. A. and Kosov, D. S. (2001), «Atom–atom partitioning of intramolecular and intermolecular Coulomb energy», *J. Chem. Phys.*, vol. 114 (15), pp. 6539–6547, URL http://aip.scitation.org/doi/10.1063/1.1356013. {Cited on page 125.}

[157] Rafat, M. and Popelier, P. L. A. (2007), «Topological Atom–Atom Partitioning of Molecular Exchange Energy and its Multipolar Convergence», in «Quantum Theory Atoms Mol.», pp. 121–140, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, URL http://doi.wiley.com/10.1002/9783527610709.ch5. {Cited on pages 121 and 122.}

[158] Reed, A. E., Curtiss, L. a. and Weinhold, F. (1988), «Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint», *Chem. Rev.*, vol. 88 (6), pp. 899–926, URL http://pubs.acs.org/doi/abs/10.1021/cr00088a005. {Cited on page 77.}

[159] Reiher, M. (2007), «On the definition of local spin in relativistic and nonrelativistic quantum chemistry.», *Faraday Discuss.*, vol. 135, pp. 97–124; discussion 125–149, 503–506. {Cited on page 100.}

[160] Resta, R. (1998), «Quantum-Mechanical Position Operator in Extended Systems», *Phys. Rev. Lett.*, vol. 80 (9), pp. 1800–1803, URL http://link.aps.org/doi/10.1103/PhysRevLett.80.1800. {Cited on page 148.}

[161] Resta, R. (2005), «Electron Localization in the Quantum Hall Regime», *Phys. Rev. Lett.*, vol. 95 (19), p. 196805, URL http://link.aps.org/doi/10.1103/PhysRevLett.95.196805. {Cited on page 148.}

[162] Resta, R. (2006), «Kohn's theory of the insulating state: A quantum-chemistry viewpoint», *J. Chem. Phys.*, vol. 124 (10), p. 104 104, URL http://aip.scitation.org/doi/10.1063/1.2176604. {Cited on pages 148 and 154.}

[163] Resta, R. and Sorella, S. (1999), «Electron Localization in the Insulating State», *Phys. Rev. Lett.*, vol. 82 (2), pp. 370–373, URL http://link.aps.org/doi/10.1103/PhysRevLett.82.370{%}5Cnhttp://arxiv.org/abs/cond-mat/9808151http://link.aps.org/doi/10.1103/PhysRevLett.82.370. {Cited on pages 148 and 196.}

[164] Santalo, N., Tarres, J., Espinosa, E., Llorca, J., Molins, E., Veciana, J., Rovira, C., Mays, M., Yang, S., Cowan, D., Garrigou-Lagrange, C., Amiell, J., Delhaes, P. and Canadell, E. (1993), «(BTDM-TTF)-TCNQ complex, a new organic metal», *Synthetic Metals*, vol. 56 (1), pp. 2050 – 2056, URL http://www.sciencedirect.com/science/article/pii/0379677993903713, proceedings of the International Conference on Science and Technology of Synthetic Metals (ICSM'92).

[165] Savin, A. (2002), «PROBABILITY DISTRIBUTIONS AND VALENCE SHELLS IN ATOMS», in «Rev. Mod. Quantum Chem.», pp. 43–62, WORLD SCIENTIFIC, URL http://www.worldscientific.com/doi/abs/10.1142/9789812775702{_}0003. {Cited on pages 66 and 78.}

[166] Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M. and Montgomery, J. A. (1993), «General atomic and molecular electronic structure system», *J. Comput. Chem.*, vol. 14 (11), pp. 1347–1363, URL http://doi.wiley.com/10.1002/jcc.540141112. {Cited on pages 104 and 154.}

[167] Shaik, S., Rzepa, H. S. and Hoffmann, R. (2013), «One Molecule, Two Atoms, Three Views, Four Bonds?», *Angew. Chemie Int. Ed.*, vol. 52 (10), pp. 3020–3033, URL http://doi.wiley.com/10.1002/anie.201208206. {Cited on page 96.}

[168] Shaik, S., Danovich, D., Wu, W., Su, P., Rzepa, H. S. and Hiberty, P. C. (2012), «Quadruple bonding in C2 and analogous eight-valence electron species», *Nat. Chem.*, vol. 4 (3), pp. 195–200, URL http://dx.doi.org/10.1038/nchem.1263http://www.nature.com/doifinder/10.1038/nchem.1263. {Cited on page 96.}

[169] Shaik, S., Danovich, D., Braida, B. and Hiberty, P. C. (2016), «A Response to a Comment by G. Frenking and M. Hermann on: "The Quadruple Bonding in C 2 Reproduces the Properties of the Molecule"», *Chem. - A Eur. J.*, vol. 22 (52), pp. 18 977–18 980, URL http://doi.wiley.com/10.1002/chem.201602840. {Cited on page 97.}

[170] Shavitt, I. (1985), «Geometry and singlet-triplet energy gap in methylene: a critical review of experimental and theoretical determinations», *Tetrahedron*, vol. 41 (8), pp. 1531–1542, URL http://linkinghub.elsevier.com/retrieve/pii/S0040402001963938. {Cited on page 112.}

[171] Sherrill, C., Van Huis, T. J., Yamaguchi, Y. and Schaefer, H. F. (1997), «Full configuration interaction benchmarks for the states of methylene», *J. Mol. Struct. THEOCHEM*, vol. 400, pp. 139–156, URL http://linkinghub.elsevier.com/retrieve/pii/S016612809790275X. {Cited on page 112.}

[172] Siegbahn, P. E. M. (1981), «Large scale contracted MC–CI calculations on acetylene and its dissociation into two CH(2Π) radicals», *J. Chem. Phys.*, vol. 75 (5), p. 2314, URL http://scitation.aip.org/content/aip/journal/jcp/75/5/10.1063/1.442294. {Cited on page 103.}

[173] Siegel, S. and Siegel, B. (1963), «Intrinsic bond energies», *J. Chem. Educ.*, vol. 40 (3), p. 143, URL http://pubs.acs.org/doi/abs/10.1021/ed040p143. {Cited on page 96.}

[174] Silvi, B. and Savin, A. (1994), «Classification of chemical bonds based on topological analysis of electron localization functions», *Nature*, vol. 371 (6499), pp. 683–686, URL http://www.nature.com/doifinder/10.1038/371683a0. {Cited on pages 121 and 147.}

[175] Slater, J. C. (1929), «The Theory of Complex Spectra», *Phys. Rev.*, vol. 34 (10), pp. 1293–1322, URL http://link.aps.org/doi/10.1103/PhysRev.34.1293. {Cited on page 24.}

[176] Souza, I., Wilkens, T. and Martin, R. M. (2000), «Polarization and localization in insulators: Generating function approach», *Phys. Rev. B*, vol. 62 (3), pp. 1666–1683, URL http://link.aps.org/doi/10.1103/PhysRevB.62.1666. {Cited on page 148.}

[177] Srebrenik, S. and Bader, R. F. W. (1975), «Towards the development of the quantum mechanics of a subspace», *J. Chem. Phys.*, vol. 63 (9), p. 3945, URL http://link.aip.org/link/JCPSA6/v63/i9/p3945/s1{&}Agg=doihttp://scitation.aip.org/content/aip/journal/jcp/63/9/10.1063/1.431834. {Cited on page 78.}

[178] Srebrenik, S., Bader, R. F. W. and Nguyen-Dang, T. T. (1978), «Subspace quantum mechanics and the variational principle», *J. Chem. Phys.*, vol. 68 (8), p. 3667, URL http://scitation.aip.org/content/aip/journal/jcp/68/8/10.1063/1.436225. {Cited on page 78.}

[179] Stalke, D., ed. (2012), *Electron Density and Chemical Bonding I: Experimental Charge Density Studies*, Electron Density and Chemical Bonding, Springer Berlin Heidelberg. {Cited on page 1.}

[180] Stalke, D., ed. (2012), *Electron Density and Chemical Bonding II: Theoretical Charge Density Studies*, Electron Density and Chemical Bonding, Springer Berlin Heidelberg. {Cited on page 1.}

[181] Stewart, R. F. (1969), «Generalized X-Ray Scattering Factors», *The Journal of Chemical Physics*, vol. 51 (10), pp. 4569–4577, URL http://scitation.aip.org/content/aip/journal/jcp/51/10/10.1063/1.1671828. {Cited on page 53.}

[182] Tang, W., Sanville, E. and Henkelman, G. (2009), «A grid-based Bader analysis algorithm without lattice bias», *J. Phys. Condens. Matter*, vol. 21 (8), p. 084 204, URL http://stacks.iop.org/0953-8984/21/i=8/a=084204?key=crossref.84a8e56b7ca1da94922dada58295eda5. {Cited on page 168.}

[183] Taraskin, S. N., Drabold, D. A. and Elliott, S. R. (2002), «Spatial Decay of the Single-Particle Density Matrix in Insulators: Analytic Results in Two and Three Dimensions», *Phys. Rev. Lett.*, vol. 88 (19), p. 196 405, URL http://link.aps.org/doi/10.1103/PhysRevLett.88.196405. {Cited on page 153.}

[184] Taraskin, S. N., Fry, P. A., Zhang, X., Drabold, D. A. and Elliott, S. R. (2002), «Spatial decay of the single-particle density matrix in tight-binding metals: Analytic results in two dimensions», *Phys. Rev. B*, vol. 66 (23), p. 233 101, URL http://link.aps.org/doi/10.1103/PhysRevB.66.233101. {Cited on page 153.}

[185] Tonner, R., Öxler, F., Neumüller, B., Petz, W. and Frenking, G. (2006), «Carbodiphosphoranes: The Chemistry of Divalent Carbon(0)», *Angew. Chemie Int. Ed.*, vol. 45 (47), pp. 8038–8042, URL http://doi.wiley.com/10.1002/anie.200602552. {Cited on page 96.}

[186] Torre, A., Alcoba, D. R., Lain, L. and Bochicchio, R. C. (2010), «Relationships between Cumulant and Spin-Density Matrices: Application to the Decomposition of Spin», *J. Phys. Chem. A*, vol. 114 (6), pp. 2344–2349, URL http://pubs.acs.org/doi/abs/10.1021/jp9090848. {Cited on page 100.}

[187] Tsirelson, V. G., Avilov, A. S., Lepeshov, G. G., Kulygin, A. K., Stahn, J., Pietsch, U. and Spence, J. C. H. (2001), «Quantitative Analysis of the Electrostatic Potential in Rock-Salt Crystals Using Accurate Electron Diffraction Data», *The Journal of Physical Chemistry B*, vol. 105 (21), pp. 5068–5074, URL http://pubs.acs.org/doi/abs/10.1021/jp0015729. {Cited on pages 2, 83, and 85.}

[188] Venkataraman, L., Klare, J. E., Nuckolls, C., Hybertsen, M. S. and Steigerwald, M. L. (2006), «Dependence of single-molecule junction conductance on molecular conformation», *Nature*, vol.

442 (7105), pp. 904–907, URL http://www.nature.com/doifinder/10.1038/nature05037. {Cited on page 146.}

[189] Waldrop, M. M. (2016), «The chips are down for Moore's law», *Nature*, vol. 530 (7589), pp. 144–147, URL http://www.nature.com/doifinder/10.1038/530144a. {Cited on page 146.}

[190] Wiberg, K. (1968), «Application of the pople-santry-segal CNDO method to the cyclopropylcarbinyl and cyclobutyl cation and to bicyclobutane», *Tetrahedron*, vol. 24 (3), pp. 1083–1096, URL http://linkinghub.elsevier.com/retrieve/pii/0040402068880573. {Cited on pages 121 and 149.}

[191] Yu, M. and Trinkle, D. R. (2011), «Accurate and efficient algorithm for Bader charge integration», *J. Chem. Phys.*, vol. 134 (6), p. 064 111, URL http://scitation.aip.org/content/aip/journal/jcp/134/6/10.1063/1.3553716. {Cited on page 168.}

[192] Zubarev, D. Y. and Boldyrev, A. I. (2008), «"Developing paradigms of chemical bonding: adaptive natural density partitioning», *Phys. Chem. Chem. Phys.*, vol. 10 (34), p. 5207, URL http://pubs.rsc.org/en/content/articlehtml/2008/cp/b804083dhttp://xlink.rsc.org/?DOI=b804083d. {Cited on page 77.}

COLOPHON

This document was typeset in LaTeX using the typographical look-and-feel `arsclassica`. Most of the graphics in this thesis are generated using `VMD`, `Inkscape` and `pgf/tikz`. The bibliography is typeset using `natbib`.