Universitat de Girona

DEVELOPMENT OF NEW TOOLS FOR LOCAL ELECTRON DISTRIBUTION ANALYSIS

Eloy Ramos Cordoba

Dipòsit legal: Gi. 832-2014
http://hdl.handle.net/10803/133376

Doctoral Thesis:

Development of New Tools for Local Electron Distribution Analysis

Eloy Ramos Cordoba

2014

Doctoral Programme in Chemistry

Supervised by:
Dr. Pedro Salvador

Presented in partial fulfillment of the requirements for a doctoral degree from the University of Girona

Dr Pedro Salvador, of University of Girona,

I declare:

That the thesis entitled "Development of New Tools for Local Electron Distribution Analysis", presented to obtain a doctoral degree, has been completed under my supervision and meets the requirements to opt for an International Doctorate.

For all intents and purposes, I hereby sign this document.

Signature

Dr. Pedro Salvador

Girona, 11 de Març de 2014

A Mireia, mi hermana.

Acknowledgments

"Ceux qui font du vélo savent que
dans la vie rien n'est jamais plat"
René Fallet

Arribats a aquest punt, només em queda agraïr molt sinceramente el suport rebut durant els últims anys. En primer lloc tinc molts motius per agraïr al meu director de tesi, en Pedro. Ya desde el final de la licenciatura estuviste dispuesto a aconsejarme en todo lo que necesité, y a cualquier hora demostrando una paciencia infinita! Hemos disfrutado, y a veces incluso sufrido, de muchas horas de ciència de las que he aprendido mucho. También hemos compartido muchos km en bicicleta y algún que otro rato de bar, que de todo se aprende. Muchas gracias por todo!

A l'Edu també li vull agraïr especialment tot el seu suport. Hem participat junts en alguns dels projectes d'aquesta tesi, i gràcies a això he pogut apendre molt. Gràcies també per la teva paciència durant el temps de descompte que he dedicat a escriure la tesi. M'agradaria també agraïr a l'István, en Martin i en Markus per la oportunitat de realitzar una estada als seus grups de recerca.

De l'IQCC en primer lloc he d'agraïr a la Carme per cuidar de mi igual que fa amb tots els membres de l'institut. A en Miquel Duran li vull agraïr l'oportunitat que em va donar per fer la tesi en aquest grup. A en Dani pels moments que tot iqcià novell ha de passar al seu costat. Gràcies a tots els membres de l'institut per crear aquest gran ambient de treball, Alexander,

Lluis, Josep Maria, Silvia, Marcel, Emili, Ramon, Sergei, Miquel Solà, etc Del sector més jove destacar la gent que ja marxava quan jo vaig començar Samat, Oscar, Mireia, David, Juanma, Quim. Amb altres hem compartit més estones com per exemple en Fala (m'has ajudat molt també amb les DMN, gràcies!!), la Sílvia (gràcies a tu i en Narcís per cuidarme a Honolulu!), en Jordi, l'Albert l'AnnaD, l'AnnaDi (moltes gràcies per la teva ajuda amb el treball de màster!), en Yevhen (compi de master), en Marc (compi canpre i de mil batallas!!), en Mauricio, en Majid (Hola amigo), en Rafa, en Juan Pablo, la Vero, en Carles, en Quansong, l'Abril, en Pep, en Mikael, la Xana, en Rambo, la Cristina, la Laia, en Sergi, en Miquel, en Quim... una bona colla d'amics!! Més enllà de l'insitut també he compartit moltes bones estones amb tota la gent del departament de química, no deixo de recordar les victòries a les JODETE!

Aviat farà deu anys que la millor promoció de químics del món es va trobar per primer cop. Tant els que heu continuat per la facultat durant aquests anys de doctorat (tigresas, tigreses, marujes, marujus, cuquis, etc), com els que heu anat marxant, moltíssimes gràcies amics meus químics tots.

No em puc deixar la gent de la Jonquera, un bonic poble del nord on hi bufa fort el vent. Segur que us sonarà a tots. Allà vaig nèixer i allà és on conservo un bon grapat d'amics de la infància. Ana, Andreu, Fernando, Manel, Menu, Miriam, Luis, etc és una sort ser el vostre amic. A la Jonquera és també on he jugat a futbol, i durant alguns anys he pogut compaginar futbol i doctorat. Un agra iment també a tots els amics que he fet gràcies a aquest esport. En especial a en Txema i en Serra, els meus mestres.

Y ya por último gracias Mama y gracias Papa. Gracias por darme la libertad para hacer siempre lo que me ha gustado. A ti Mireia te dedico la tesis, que mas voy a decir! Gracias a toda mi familia, a los que estais mas cerca, Tio, Tia, Eze, Cynthia, Core, Lucia, Isas, Antonia y familia. Gracias tambien a los que estais mas lejos, me gustaría que nos puediéramos ver mas a menudo. Y también gracias a los que ya no están, en especial a mis dos abuelas por haber cuidado siempre de mi, Francisca y Julia.

Full list of Publications

This Thesis is presented as a compendium of publications.

Published articles included in this Thesis:

1. Ramos-Cordoba, E.; Matito, E.; Mayer, I.; Salvador, P. **Toward a Unique Definition of the Local Spin** *J. Chem. Theory Comput.* **2012**, *8*, 1270–1279. (Impact factor: 5.389, position 3/34 in PHYSICS, ATOMIC, MOLECULAR & CHEMICAL, 1^{st} quartile)

2. Ramos-Cordoba, E.; Matito, E.; Salvador, P.; Mayer, I. **Local Spins: Improved Hilbert-Space Analysis** *Phys. Chem. Chem. Phys.* **2012**, *14*, 15291–15298. (Impact factor: 3.829, position 6/34 in PHYSICS, ATOMIC, MOLECULAR & CHEMICAL, 1^{st} quartile)

3. Ponec, R.; Ramos-Cordoba, E.; Salvador, P. **Bonding Quandary in the $[Cu_3S_2]^{3+}$ Core: Insights from the Analysis of Domain Averaged Fermi Holes and the Local Spin** *J. Phys. Chem. A* **2013**, *117*, 1975–1982. (2011 Impact factor: 2.771, position 10/34 in PHYSICS, ATOMIC, MOLECULAR & CHEMICAL, 2^{nd} quartile)

4. Ramos-Cordoba, E.; Salvador, P.; Mayer, I. **The Atomic Orbitals of the Topological Atom** *J. Chem. Phys.* **2013**, *138*, 214107–214116. (Impact factor: 3.164, position 8/34 in PHYSICS, ATOMIC, MOLECULAR & CHEMICAL, 1^{st} quartile)

5. Ramos-Cordoba, E.; Salvador, P.; Reiher, M. **Local Spin and Chemical Bonding** *Chem. Eur. J.* **2013**, *19*, 15267–15275. (Impact factor: 5.831, position 21/152 in CHEMISTRY, MULTIDISCIPLINARY, 1^{st} quartile)

6. Salvador, P.; Ramos-Cordoba, E. **An Approximation to Bader's Topological Atom** *J. Chem. Phys.* **2013**, *139*, 071103. (Impact factor: 3.164, position 8/34 in PHYSICS, ATOMIC, MOLECULAR & CHEMICAL, 1^{st} quartile)

7. Ramos-Cordoba, E.; Salvador, P. **Characterization and Quantification of Polyradical Character** *J. Chem. Theory Comput.* **2014**, *10*, 634–641. (Impact factor: 5.389, position 3/34 in PHYSICS, ATOMIC, MOLECULAR & CHEMICAL, 1^{st} quartile)

Submitted articles included in this Thesis:

8. Ramos-Cordoba, E.; Salvador, P. **Diradical Character from the Local Spin Analysis**, *submitted*, **2014**

9. Ramos-Cordoba, E.; Salvador, P. **Oxidation States from Wave Function Analysis**, *submitted*, **2014**

List of Abbreviations

Abbreviation	Description
1-RDM	First-order density matrix
2-RDM	Second-order density matrix
AIM	Atom in molecule
BO	Bond order
CASSCF	Complete active space self consistent field
CP	Critical point
DAFH	Domain average Fermi holes
DFT	Density functional theory
DI	Delocalization index
eff-AO	Effective atomic orbital
EOS	Effective oxidation state
FCI	Full configuration interaction
LCAO	Linear combination of atomic orbitals
MO	Molecular orbital
OS	Oxidation state
QTAIM	Quantum theory of atoms in molecules
ROHF	Restricted open-shell Hartree Fock
SVD	Singular value decomposition
TFVC	Topological fuzzy Voronoi cell

List of Figures

List of Tables

Contents

Summary of the Thesis

The identification of an atom within the molecule is a crucial concept that permits the decomposition of physical quantities into atomic (and diatomic) contributions. Bader's Quantum Theory of Atoms in Molecules (QTAIM) still provides nowadays one of the most widely used and accepted definitions for an atom in the molecule. In QTAIM, the topology of the electron density is cleverly used to derive the atomic basins and the sharp boundaries that define them. There are two disadvantages to the method, namely the appearance of (spurious) non-nuclear attractors (NNA), which makes the analysis in chemical terms rather difficult, and the computational cost associated to the construction of the complex-shaped atomic domains.

In this thesis a new topological definition of fuzzy Voronoi cells is proposed, reproducing to some extent the results of the QTAIM analysis at a much reduced computational cost. The new atomic definition is also flexible enough to either ignore the presence of spurious non-nuclear attractors or to readily incorporate them in the analysis.

Clearly, the mere subdivision of the physical space into atomic basins does not render any atomic property other than their volume. Instead, we are interested in assigning different physical quantities to the individual atoms (or molecular fragments), which will be associated to the effective contributions of that quantity within the molecule. A second key ingredient to extract chemical information from wave function analysis is thus the proper definition of the physical quantities of interest, typically in terms of the integral over the space of the appropriate one- (two) particle density functions. Two main factors are focused in this thesis: the local spin and the effective orbitals of the atom within the molecule.

The spin properties of a system are typically characterized by the spin density (difference between alpha and beta electron densities). However, there are singlet systems such as diradicals or antiferromagnets for which the spin density vanishes in every point of the space, but still one assumes the existence of some localized spins. Several formulations have been proposed in the literature to account for the presence of local spins in molecular systems from ab initio wave functions.

In this thesis we have derived a new general formulation based on the decomposition of the expectation value of the spin squared operator into atomic and diatomic contributions, in such a way that the atomic terms reveal the existence of effective local spin in the atoms or fragments, and the diatomic ones inform about the existence and nature of the couplings between the local spins. This new formulation fulfills all necessary physical requirements that have been imposed to date. In particular, it is able to distinguish the situations of two spins coupled to form a bond from a genuine antiferromagnetic interaction. Appropriate formulations have been derived for 3D-space atomic definitions, as well as for the so-called Hilbert-space analysis. The latter proved to be a non-trivial task due to an inherent ambiguity in its application to the decomposition of genuine two-electron quantities.

The examination of the one- and two-center contributions is what we refer to the local spin analysis. In this thesis we have applied this new methodology to a number of chemical problems of interest. In particular, we have analyzed the connection between the local spin analysis and the nature of the chemical bond. It has been found that the presence of significant local spins in bonded molecules flags deviations from the classical bonding prototypes, that is, the extent of local spin is a measure of the deviation of a bond from perfect covalent character. A remarkable finding of this work is that the C_2 molecule in its ground state has all ingredients to be categorized as a diradical, ruling out the existence of a conventional quadruple bond.

We have also applied the local spin analysis to detect and quantify the polyradical character in molecules. The descriptors obtained from the local spin analysis can be used to define a general index, $\Delta^{(k)}$, that measures how close a real molecular system is to an ideal system of k perfectly localized spin centers.

The main advantage of this index is that, contrary to most approaches in the literature, it is calculated in the same manner from both multireference and unrestricted single-determinant wave functions, for any electronic state.

In the case of diradicals and diradicaloids, the $\Delta^{(0)}$ and $\Delta^{(2)}$ indexes have been used to measure deviation from a nonradical and a perfect diradical picture, respectively. The first attempt of actual quantification of triradical character has been accomplished using the values of the $\Delta^{(3)}$ index. The electronic structure of prototype radical systems such as benzyne isomers, propellanes and a number of organic diradical and triradicals have been scrutinized in detail to illustrate the new approach.

Another tool that aims at the characterization of the state of the atom within the molecule is Mayer's effective atomic orbitals (eff-AOs). In this approach the net atomic population is expressed in terms of an orthonormal set of atomic hybrids and their occupation numbers. The shape and occupation number of the hybrids faithfully reproduce the core and valence shells of the atoms.

In this thesis we have carried out for the first time the eff-AOs in the framework of QTAIM analysis for a general wave function. Apart from their conceptual significance, we have shown that the molecular orbitals of a calculation can be exactly expressed as a linear combination of this orthonormalized set of numerical atomic orbitals. This has also permitted to establish a connection between Hilbert-space and 3D-space analysis: Mulliken population analysis carried out on this basis set exactly reproduces the original QTAIM atomic populations of the atoms.

On the other hand, one cannot withstand the importance of the classical concept of oxidation state (OS). In coordination chemistry, the OS of the metal atom is typically defined as "the charge it would bear if all the ligands were removed along with the electron pairs that were shared with the central atom". Thus, formal OS are obtained by assigning integer number of electrons to the atoms/ligands according to some rules, which however, in complicated bonding situations, may be rather ambiguous. Oxidation states are intrinsically related to the electronic distribution, but electronic or spin populations are only a pointer of the atom's OS.

In this thesis we have shown that the eff-AOs can be utilized to derive the

most appropriate atomic electron configurations for the atoms or molecular fragments in the molecule. Accordingly, we have introduced a simple and general way to derive effective oxidation states (EOS) from the analysis of the electron density. Moreover, the occupation numbers of the eff-AOs can be used to quantify to which extent a particular wave function conforms to the ideal OS assignation. The method is general and it can formally be applied to any molecular system but the focus is on organometallic complexes for which the knowledge of the OS is important to understand the chemistry of the complex. For practical applications on transition metal complexes we have proposed a hierarchical strategy based on the partitioning of the molecular system into fragment/ligands prior to the eff-AO analysis.

Finally, the combination of local spin analysis and the eff-AOs (among other tools) have permitted us to gain a deeper insight into the elusive nature of the $[Cu_3S_2]^{3+}$ core. In particular, the scrutiny of the eff-AOs points to a predominant formal Cu^{II} oxidation state, whereas the lack of any significant local spins in the core is at odds with the suggested existence of antiferromagnetic coupling among the Cu and S atoms.

Resum de la Tesi

La identificació d'un àtom dins d'una molècula és un concepte fonamental que permet la descomposició de quantitats físiques en contribucions atòmiques (i diatòmiques) . La *teoria d'àtoms en molècules de Bader* (QTAIM) proporciona avui en dia una de les definicions més àmpliament utilitzades i acceptades per a un àtom en la molècula. D'acord amb la QTAIM, la topologia de la densitat electrònica s'utilitza per trobar les conques atòmiques i els límits estrictes que les defineixen. El mètode presenta dos desavantatges: per una banda, l'aparició dels anomenats *atractors no nuclears* (NNA), que dificulta l'anàlisi en termes químics, i per altra banda el cost computacional associat a la construcció dels dominis atòmics. En aquesta tesi es proposa una nova definició de les *cel.les de Voronoi difuses* que reprodueixen, en certa mesura, els resultats de l'anàlisi QTAIM amb un cost computacional molt menor. La nova definició atòmica permet ignorar la presència de NNA o incorporar-los fàcilment en l'anàlisi.

És evident que la subdivisió de l'espai físic en conques atòmiques no representa cap propietat atòmica en sí mateixa a part del volum àtomic. El que resulta interessant és l'assignació de diferents magnituds físiques locals als àtoms o fragments moleculars. Deixant de banda la definició atòmica, un ingredient clau per extreure informació química a partir de l'anàlisis de la funció d'ona és la definició adequada de les quantitats físiques, típicament en termes de la integral sobre l'espai de les funcions densitats corresponents a una o dues partícules. En aquesta tesi ens hem centrat en dues propietats: L'espín local (*local spin*) i els orbitals àtomics efectius (*effective atomic orbitals*).

Les propietats d'espín d'un sistema es caracteritzen típicament per la *densitat d'espín*(diferència entre les densitats electròniques alfa i beta). No obstant,

hi ha sistemes singlet com ara dirradicals o complexes antiferromagnètics pels que la densitat d'espín s'esvaeix en cada punt de l'espai, tot i que s'assumeix l'existència d'espins locals. S'han proposat diverses formulacions en la literatura per donar compte de la presència d'espins locals en sistemes moleculars a partir de les funcions d'ona *ab initio*. En aquesta tesi hem derivat una nova formulació general basada en la descomposició del valor esperat de $\langle \hat{S}^2 \rangle$, en contribucions atòmiques i diatòmiques, de tal manera que els termes atòmics revelen l'existència d'espins locals en els àtoms o fragments i les diatòmiques informen sobre l'existència i la naturalesa dels acoblaments entre ells. Aquesta nova formulació compleix amb tots els requisits físics necessaris que s'han imposat fins ara. En particular, és capaç de distingir entre dos espins acoblats per formar un enllaç covalent i una veritable interacció antiferromagnética.

L'anàlisi de les contribucions d'un i dos centres a $\langle \hat{S}^2 \rangle$ és el que anomenem anàlisi de l'espín local. En aquesta tesi hem aplicat aquesta nova metodologia a una sèrie de problemes químics d'interès. En particular, s'ha analitzat la relació entre l'espín local i la naturalesa de l'enllaç químic. S'ha trobat que la presència d'espín local indica desviacions dels prototips d'enllaç clàssics, és a dir, el grau d'espín local és una mesura de la desviació del caràcter covalent d'un enllaç. Una troballa notable d'aquest treball és que la molècula C_2 en el seu estat fonamental té tots els ingredients per ser classificada com un diradical, descartant l'existència d'un quart enllaç convencional.

També hem aplicat el *local spin analysis* per detectar i quantificar el caràcter poliradical en molècules. Els valors d'espín local s'han utitlitzat per definir un nou índex, $\Delta^{(k)}$, que estima la desviació d'un sistema molecular real respecte un sistema ideal de k espins perfectament localitzats. El principal avantatge d'aquest índex és que, contràriament a la majoria dels descriptors de la bibliografia, es pot utilitzar tant per funcions d'ona multireferencials com per funcions d'ona d'un sol determinant d'Slater, així com per a qualsevol estat electrònic.

En el cas de dirradicals i diradicaloids, $\Delta^{(0)}$ i $\Delta^{(2)}$ s'han utilitzat per a mesurar la desviació respecte d'un no radical i un sistema diradical perfecte respectivament. El primer intent de quantificació real del caràcter triradical s'ha aconseguit mitjançant valors de l'índex $\Delta^{(3)}$. L'estructura electrònica de diradicals

i triradicals orgànics s'ha analitzat en detall per il.lustrar el nou mètode.

Una altra eina que té com a objectiu la caracterització de l'estat de l'àtom dins de la molècula són els *orbitals àtomics efectius* (eff-AOs) de Mayer. En aquest context la població atòmica neta s'expressa en termes d'un conjunt ortonormal d'híbrids atòmics i els seus números d'ocupació. La forma i les ocupacions dels híbrids reprodueixen fidelment els orbitals interns i els de valència dels àtoms. En aquesta tesi hem implementat per primera vegada els eff-AOs en el marc de l'anàlisi QTAIM per una funció d'ona general. A més de la seva importància conceptual, hem demostrat que els orbitals moleculars d'un càlcul es poden expressar exactament com una combinació lineal d'aquest conjunt ortonormalizat d'orbitals atòmics numèrics. Això també ha permès establir una connexió entre els anàlisis en els espais de Hilbert i l'espai tridimensional.

D'altra banda, el concepte clàssic d'estat d'oxidació (OS) és un dels més importants dins l'àmbit de la química. En la química de coordinació, l'OS es defineix generalment com "la càrrega que queda al metall després de retirar tots els lligands en la seva configuració de capa tancada". Per tant, els OS formals s'obtenen mitjançant l'assignació de nombres enters d'electrons als àtoms / lligands d'acord amb algunes regles predefinides. No obstant, hi ha situacions en que aquesta assignació pot ser bastant ambigua. Els OS estàn intrínsecament relacionats amb la distribució electrònica, però les poblacions electròniques o d'espín són només una primera indicació de l'OS de l' àtom.

En aquesta tesi s'ha demostrat que els eff-AOs es poden utilitzar per derivar les configuracions electròniques més apropiades per als àtoms o fragments moleculars en la molècula . En conseqüència, hem introduït una forma simple i general per derivar estats d'oxidació eficaços (EOS) a partir de l'anàlisi de la densitat electrònica. D'altra banda, els números d'ocupació dels eff-AOs es poden utilitzar per quantificar en quina mesura una funció d'ona en particular s'ajusta a l'assignació ideal d'OS. El mètode és general i formalment es pot aplicar a qualsevol sistema molecular pero en aquest treball ens centrem en complexes organometàl.lics, que són els més rellevants en aquest context. Com a aplicació pràctica, en els complexos de metalls de transició hem proposat una estratègia jeràrquica basada en la partició del sistema molecular en fragments/lligands abans de l'anàlisi dels eff-AOs.

Finalment, la combinació de l'anàlisi de l'espín local i els eff-AOs (entre altres eines) ens han permès obtenir una visió més ajustada de la naturalesa del complexe $[Cu_3S_2]^{3+}$. En particular, els eff-AOs indiquen un estat d'oxidació formal Cu^{II}, mentre que la manca d'espins locals significatius en el nucli va en contra de l'existència d'acoblaments antiferromagnètics entre els àtoms de Cu i S.

Resumen de la Tesis

La identificación de un átomo dentro de una molécula es un concepto fundamental que permite la descomposición de cantidades físicas en contribuciones atómicas (y diatómicas). La teoría de àtomos en moléculas de Bader (QTAIM) proporciona hoy en día una de las definiciones más ampliamente utilizadas y aceptadas para un átomo en la molécula. De acuerdo con la QTAIM, la topología de la densidad electrónica se utiliza para encontrar las cuencas atómicas y los límites estrictos que las definen. El método presenta dos desventajas: por un lado, la aparición de los llamados atractores no nucleares (NNA), que dificulta el análisis en términos químicos, y por otro lado el coste computacional asociado a la construcción de los dominios atómicos. En esta tesis se propone una nueva definición de las celdas de Voronoi difusas que reproducen, en cierta medida, los resultados del análisis QTAIM con un coste computacional muy menor. La nueva definición atómica permite ignorar la presencia de NNA o incorporarlos fácilmente en el análisis.

Es evidente que la subdivisión del espacio físico en cuencas atómicas no representa ninguna propiedad atómica en sí misma mas que la del volumen atómico. Lo que resulta interesante es la asignación de diferentes magnitudes físicas locales a los átomos o fragmentos moleculares. Dejando de lado la definición atómica, un ingrediente clave para extraer información química a partir del análisis de la función de onda es la definición adecuada de las cantidades físicas, típicamente en términos de la integral sobre el espacio de las funciones densidades correspondientes a una o dos partículas. En esta tesis nos hemos centrado en dos propiedades: El espín local (*local spin*) y los orbitales atómicos efectivos (*effective atomic orbitals*).

Las propiedades de espín de un sistema se caracterizan típicamente por la

densidad de espín (diferencia entre las densidades electrónicas alfa y beta). No obstante, hay sistemas singulete como por ejemplo dirradicales o complejos antiferromagnètics por los que la densidad de espín se desvanece en cada punto del espacio, a pesar de que se asume la existencia de espines locales. Se han propuesto varias formulaciones en la literatura para dar cuenta de la presencia de espines locales en sistemas moleculares a partir de las funciones de onda *ab initio*. En esta tesis hemos derivado una nueva formulación general basada en la descomposición del valor esperado de $\langle \hat{S}^2 \rangle$, en contribuciones atómicas y diatómicas, de tal manera que los términos atómicos revelan la existencia de espines locales en los átomos o fragmentos y las diatómicas informan sobre la existencia y la naturaleza de los acoplamientos entre ellos. Esta nueva formulación cumple con todos los requisitos físicos necesarios que se han impuesto hasta ahora. En particular, es capaz de distinguir entre dos espines acoplados para formar un enlace covalente y una verdadera interacción antiferromagnética.

El análisis de las contribuciones de uno y dos de centros a $\langle \hat{S}^2 \rangle$ es lo que denominamos anàlisis del espín local. En esta tesis hemos aplicado esta nueva metodología a una serie de problemas químicos de interés. En particular, se ha analizado la relación entre el espín local y la naturaleza del enlace químico. Se ha encontrado que la presencia de espín local indica desviaciones de los prototipos de enlace clásicos, el grado de espín local es una medida de la desviación del carácter covalente de un enlace. Un hallazgo notable de este trabajo es que la molécula C_2 en su estado fundamental tiene todos los ingredientes para ser clasificada como un diradical, descartando la existencia de un cuarto enlace convencional. También hemos aplicado el local spin analysis para detectar y cuantificar el carácter poliradical en moléculas. Los valores de espín local se han utitlizado para definir un nuevo índice, $\Delta^{(k)}$, que estima la desviación de un sistema molecular real frente a un sistema ideal de k espines perfectamente localizados. La principal ventaja de este índice es que, contrariamente a la mayoría de los descriptores de la bibliografía, se puede utilizar tanto por funciones de onda multireferenciales cómo para funciones de onda de un solo determinante de Slater, así como para cualquier estado electrónico. En el caso de dirradicales y diradicaloides, $\Delta^{(0)}$ y $\Delta^{(2)}$ se han utilizado para medir la

desviación respecto de un no radical y un sistema diradical perfecto respectivamente. El primer intento de cuantificación real del carácter triradical se ha conseguido mediante valores del índice $\Delta^{(3)}$. La estructura electrónica de diradicales y triradicales orgánicos se ha analizado en detalle para ilustrar el nuevo método.

Otra herramienta que tiene como objetivo la caracterización del estado del átomo dentro de la molécula son los orbitales atómicos efectivos (eff-AOs) de Mayer. En este contexto la población atómica neta se expresa en términos de un conjunto ortonormal de híbridos atómicos y sus números de ocupación. La forma y las ocupaciones de los híbridos reproducen fielmente los orbitales internos y los de valéncia de los átomos. En esta tesis hemos implementado por primera vez los eff-AOs en el marco del análisis QTAIM para una función de onda general. Además de su importancia conceptual, hemos demostrado que los orbitales moleculares de un cálculo se pueden expresar exactamente como una combinación lineal de este conjunto ortonormalizado de orbitales atómicos numéricos. Esto también ha permitido establecer una conexión entre los análisis en el espacios de Hilbert y el espacio tridimensional. Por otro lado, el concepto clásico de estado de oxidación (OS) es uno de los más importantes dentro del ámbito de la química. En la química de coordinación, el OS se define generalmente como 'la carga que queda al metal después de retirar todos los ligandos en su configuración de capa cerrada'. Por lo tanto, los OS formales se obtienen mediante la asignación de números enteros de electrones a los átomos/ligandos de acuerdo con algunas reglas predefinidas. No obstante, hay situaciones en que esta asignación puede ser bastante ambigua. Los OS están intrínsecamente relacionados con la distribución electrónica, pero las poblaciones electrónicas o de espín son sólo una primera indicación del OS del átomo.

En esta tesis se ha demostrado que los eff-AOs se pueden utilizar para derivar las configuraciones electrónicas más apropiadas para los átomos o fragmentos moleculares en la molécula. En consecuencia, hemos introducido una forma simple y general para derivar estados de oxidación eficaces (EOS) a partir del análisis de la densidad electrónica. Por otro lado, los números de ocupación de los eff-AOs se pueden utilizar para cuantificar hasta que punto una función

de onda en particular se ajusta a la asignación ideal de OS. El método es general y formalmente se puede aplicar a cualquier sistema molecular pero en este trabajo nos centramos en complejos organometàlicos, que son los más relevantes en este contexto. Como aplicación práctica, en los complejos de metales de transición hemos propuesto una estrategia jerárquica basada en la partición del sistema molecular en fragmentos/ligandos antes del análisis de los eff-AOs.

Finalmente, la combinación del análisis del espín local y los eff-AOs (entre otras herramientas) nos han permitido obtener una visión más ajustada de la naturaleza del complejo $[Cu_3S_2]^{3+}$. En particular, los eff-AOs indican un estado de oxidación formal Cu^{II}, mientras que la carencia de espines locales significativos en el núcleo va en contra de la existencia de acoplamientos antiferromagnèticos entre los átomos de Cu y S.

Chapter 1

Introduction

The growth in computing power over the last decades alongside with the advances in electronic structure methods have increased the reliability of predicting molecular properties. Unfortunately, the more sophisticated the electronic structure are, the further those methods are to the language of chemistry.[1]

Quantum chemistry is written in the language of physics, that is, a wave function describing particles and their interactions. Therefore, a molecule is viewed as a system of electrons and nuclei interacting with each other. From a classical point of view, chemistry deals with atoms that are kept together by bonds to form a molecule. This view is still necessary and useful for the prediction of chemical properties and for the elucidation of chemical structures.[2] Quantum mechanics puts forward the Schrödinger equation whose solution leads to the exact wave function. However, in practice this equation is always solved approximately. It is noteworthy to remark that some of the approximations made to solve the Schrödinger equation are based on classical or pre-quantum chemistry concepts. For example, the valence bond method is based on the electron-pair bond formalism.[3]

While some researchers advocate for the abandon of the chemical view and several concepts accompanying it (bond order, bond ionicity or aromaticity) this jargon is deeply rooted in the chemistry community and it cannot be so easily ignored. It thus is of high importance to establish bridges between the

two views of a molecular system: the "physical" one of interacting particles and the "chemical" one of atoms kept together by bonds. This thesis aims to shed some light into the field of extracting chemical information from wave function analysis. The thesis can be divided into three parts. The first one deals with the formulation of a new atom in molecule definition. In the second part we propose a new methodology to obtain local spins from wave function analysis. Finally, we study the electron configurations of the atom within the molecule and retrieve their oxidation states from a particular analysis of the effective atomic orbitals (eff-AOs).

1.1 Wave functions, reduced density matrices and cumulants

A pure quantum state can be fully described by a wave function $\Psi_i(\vec{x}_1, ..., \vec{x}_N)$, where N is the number of electrons and $\vec{x}$ their spatial and spin coordinates ($\vec{x} = \vec{r}, \vec{\sigma}$). Within the Born-Oppenheimer approximation, the wave functions are typically obtained as approximate solutions of the time-independent non-relativistic electronic Schrödinger equation

$$\hat{H}\Psi_i(\vec{x}_1, ..., \vec{x}_N) = E_i\Psi_i(\vec{x}_1, ..., \vec{x}_N) \tag{1.1}$$

where $\hat{H}$ is the electronic Hamiltonian of the system that only depends on the spatial coordinates of the particles, and E_i are the energies of the different states. Moreover, according to the antisymmetry principle, the wave function must be antisymmetric with respect to the interchange of coordinates of two particles, for instance

$$\Psi(\vec{x}_1, ..., \vec{x}_i, ..., \vec{x}_j, ..., \vec{x}_N) = -\Psi(\vec{x}_1, ..., \vec{x}_j, ..., \vec{x}_i, ..., \vec{x}_N). \tag{1.2}$$

In the case of mixed or ensembles states, the description of the system is provided by the so-called N-density matrices,[4]

$$\rho_N(\vec{x}'_1, ..., \vec{x}'_N; \vec{x}_1, ..., \vec{x}_N) = \Psi^*(\vec{x}'_1, ..., \vec{x}'_N)\Psi(\vec{x}_1, ..., \vec{x}_N). \tag{1.3}$$

In this sense, the N-density matrix is a more complete object than the wave function. The diagonal elements ($\vec{x}' = \vec{x}$) of 1.3 are related to the probability density or probability distribution of the system according to the Born interpretation of the wave function. The product

$$\rho_N(\vec{x}_1, ..., \vec{x}_N; \vec{x}_1, ..., \vec{x}_N)d\vec{x}_1 \cdots d\vec{x}_N \tag{1.4}$$

gives us the probability of finding simultaneously electron 1 between $\vec{x}_1$ and $\vec{x}_1 + d\vec{x}_1$, electron 2 between $\vec{x}_2$ and $\vec{x}_2 + d\vec{x}_2$ and so on.

From the N-density matrix we can obtain lower-rank density matrices by integration of the appropriate coordinates. Of particular interest for the purpose of this thesis are the first- and second-order reduced density matrices (1-RDM and 2-RDM, respectively). The operators in quantum chemistry only involves the coordinates of one- or two-electron, therefore, their expectation values can be fully written in terms of the first- and second- order RDM. The 1-RDM is defined as

$$\rho(\vec{x}'_1; \vec{x}_1) = N \int \rho_N(\vec{x}'_1, \vec{x}_2, ..., \vec{x}_N; \vec{x}_1, \vec{x}_2, ..., \vec{x}_N)d\vec{x}_2 \cdots d\vec{x}_N. \tag{1.5}$$

Its diagonal elements are the electron density $\rho(\vec{x})$, that is related to the probability of finding one electron between $\vec{x}$ and $\vec{x}+d\vec{x}$ independently of the position of the other electrons. By integrating the spin coordinates of the diagonal part of eq. (1.4) one obtains the so-called spinless density

$$\rho(\vec{r}_1) = \rho^\alpha(\vec{r}_1) + \rho^\beta(\vec{r}_1), \tag{1.6}$$

which depends only on the spatial coordinates of the electrons ($\vec{r}$). Simultaneously, one can define the spin density as

$$\rho^s(\vec{r}_1) = \rho^\alpha(\vec{r}_1) - \rho^\beta(\vec{r}_1). \tag{1.7}$$

The 2-RDM, assuming the so-called McWeeny normalization, reads as

$$\rho_{(2)}(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2) = N(N-1) \int \rho_N(\vec{x}_1', \vec{x}_2, ..., \vec{x}_N; \vec{x}_1, \vec{x}_2, ..., \vec{x}_N) d\vec{x}_3 \cdots d\vec{x}_N \tag{1.8}$$

It is Hermitian

$$\rho_{(2)}(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2) = \rho^*_{(2)}(\vec{x}_1, \vec{x}_2; \vec{x}_1', \vec{x}_2'), \tag{1.9}$$

antisymmetric w.r.t. each set of indexes

$$\rho_{(2)}(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2) = -\rho_{(2)}(\vec{x}_2', \vec{x}_1'; \vec{x}_1, \vec{x}_2) \tag{1.10}$$

$$\rho_{(2)}(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2) = -\rho_{(2)}(\vec{x}_1', \vec{x}_2'; \vec{x}_2, \vec{x}_1), \tag{1.11}$$

and symmetric with respect to particle permutation

$$\rho_{(2)}(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2) = \rho_{(2)}(\vec{x}_2', \vec{x}_1'; \vec{x}_2, \vec{x}_1) \tag{1.12}$$

By integrating over the spin coordinates

$$\rho_{(2)}(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) = \int \rho_{(2)}(\vec{x}_1', \vec{x}_2'; \vec{x}_1, \vec{x}_2)|_{\vec{\sigma}_1' = \vec{\sigma}_1, \vec{\sigma}_2' = \vec{\sigma}_2} d\vec{\sigma}_1 d\vec{\sigma}_2, \tag{1.13}$$

we can obtain the so-called spinless 2-RDM as a sum of the spin components as

$$\begin{aligned} \rho_{(2)}(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) &= \rho^{\alpha\alpha\alpha\alpha}_{(2)}(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) + \rho^{\beta\beta\beta\beta}_{(2)}(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) \\ &\quad + \rho^{\alpha\beta\alpha\beta}_{(2)}(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) + \rho^{\beta\alpha\beta\alpha}_{(2)}(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2). \end{aligned} \tag{1.14}$$

The spinless 2-RDM can be written in terms of the the spinfree cumulant, $\Gamma(\vec{r}_1, \vec{r}_2; \vec{r}_1', \vec{r}_2')$, the genuine two-electron quantity, and the spinless 1-RDM:

$$\begin{aligned} \rho_2(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) &= \rho(\vec{r}_1'; \vec{r}_1)\rho(\vec{r}_2'; \vec{r}_2) - \tfrac{1}{2}\rho(\vec{r}_1'; \vec{r}_2)\rho(\vec{r}_2'; \vec{r}_1) \\ &\quad - \tfrac{1}{2}\rho^s(\vec{r}_1'; \vec{r}_2)\rho^s(\vec{r}_2'; \vec{r}_1) + \Gamma(\vec{r}_1', \vec{r}_2'; \vec{r}_1, \vec{r}_2) \end{aligned} \tag{1.15}$$

For single-determinant wave functions, the 2-RDM can be entirely written in terms of the 1-RDM because their cumulant matrix is zero. The diagonal of eq. (1.14) is the so-called pair density, $\rho(\vec{r}_1, \vec{r}_2)$. This density has been extensively

employed to study the electron pair density structure of the chemical bond.[5] Another widely used density used in chemical bonding analysis is the exchange-correlation density[6]

$$\rho_{XC}(\vec{r}_1, \vec{r}_2) = \rho(\vec{r}_1, \vec{r}_2) - \frac{1}{2}\rho(\vec{r}_1)\rho(\vec{r}_2) \tag{1.16}$$

We will take advantage of the matrices and densities defined in this section to develop tools to describe electron localization in molecules. For practical applications, we will use the program DMN,[7] written by Matito and Feixas, in order to construct 1-RDM and 2-RDM from commercial quantum chemistry packages, such as Gaussian03.[8]

1.2 Electron spin

In 1924, Pauli demonstrated that four quantum numbers are necessary to characterize the electrons, thus explaining the shell-like structure of an atom. Actually the contribution of Pauli was the fourth quantum number itself, the spin, an intrinsic property of the electron as proved a year later in the experiment of Uhlenbeck and Goudsmit. In quantum mechanics, spin represents an angular momentum. The spin is a pure quantum property, it does not have any classical analogue. As an angular momentum the spin of an electron is described by a vector, for which each of its components is associated with the corresponding operator

$$\hat{s} = (\hat{s}_x, \hat{s}_y, \hat{s}_z). \tag{1.17}$$

From these one can define the operator associated to the modulus of the spin angular momentum (total spin squared operator)

$$\hat{s}^2 = \hat{s}_x^{\,2} + \hat{s}_y^{\,2} + \hat{s}_z^{\,2}. \tag{1.18}$$

From the postulates of quantum mechanics it can be derived that the operators of any angular momentum motion (it also applies to orbital angular momentum) fulfill the following commutation rules (in a.u.)

$$[\hat{s}_x, \hat{s}_y] = i\hat{s}_z, \quad [\hat{s}_x, \hat{s}_z] = i\hat{s}_y, \quad [\hat{s}_y, \hat{s}_z] = i\hat{s}_x$$
$$[\hat{s}^2, \hat{s}_x] = [\hat{s}^2, \hat{s}_y] = [\hat{s}^2, \hat{s}_z] = 0 \tag{1.19}$$

These conditions prevent the full determination, in general, of the three components of the spin vector. What can be known is the modulus and one of the components (usually the z component is chosen). It is easy to show that these quantities are governed by a single parameter, the spin quantum number s. The spin functions $f(\omega)$, where ω represents the spin coordinate, are eigenfunction of the $\hat{s}^2$ operator

$$\hat{s}^2 f(\omega) = s(s+1)f(\omega), \tag{1.20}$$

and the $\hat{s}_z$

$$\hat{s_z} f(\omega) = m_s f(\omega) \tag{1.21}$$

where m_s is the projection of the spin vector on a chosen axis (the z-component). Since electrons have s=$\frac{1}{2}$, and m_s=[-$\frac{1}{2}$, $\frac{1}{2}$], the spin states of an electron can be described by two functions, $\alpha(\omega)$ (s=1/2 and m_s=1/2) and $\beta(\omega)$ (s=1/2 and m_s=-1/2). These spin functions are orthonormal

$$\int \alpha^*(\omega)\alpha(\omega)d\omega = \int \beta^*(\omega)\beta(\omega)d\omega = 1, \tag{1.22}$$

and

$$\int \alpha^*(\omega)\beta(\omega)d\omega = \int \beta^*(\omega)\alpha(\omega)d\omega = 0. \tag{1.23}$$

The action of the different individual one-electron spin operators over the corresponding alpha and beta one-electron spin functions yields the well-known relations

$$\hat{s}_x\alpha(\omega) = \frac{1}{2}\beta(\omega) \quad \hat{s}_y\alpha(\omega) = \tfrac{i}{2}\beta(\omega) \quad \hat{s}_z\alpha(\omega) = \frac{1}{2}\alpha(\omega)$$
$$\hat{s}_x\beta(\omega) = \frac{1}{2}\alpha(\omega) \quad \hat{s}_y\beta(\omega) = -\tfrac{i}{2}\alpha(\omega) \quad \hat{s}_z\beta(\omega) = -\frac{1}{2}\beta(\omega) \tag{1.24}$$

and

$$\hat{s}^2\alpha(\omega) = \frac{3}{4}\alpha(\omega) \qquad \hat{s}^2\beta(\omega) = \frac{3}{4}\beta(\omega). \tag{1.25}$$

For an N-electron system, the total spin operator is given by the components

$$\hat{S}_x = \sum_i^N \hat{s}_x(i) \quad \hat{S}_y = \sum_i^N \hat{s}_y(i) \quad \hat{S}_z = \sum_i^N \hat{s}_z(i) \tag{1.26}$$

and its modulus

$$\begin{aligned}\hat{S}^2 = \sum_i^N |\vec{s}_i|^2 + \sum_{i\neq j}^N \vec{s}_i\vec{s}_j = \sum_i^N \hat{s}_x^2(i) + \hat{s}_y^2(i) + \hat{s}_z^2(i) + \\ + \sum_{i\neq j}^N \hat{s}_x(i)\hat{s}_x(j) + \hat{s}_y(i)\hat{s}_y(j) + \hat{s}_z(i)\hat{s}_z(j).\end{aligned} \tag{1.27}$$

From eq. (1.27) it can be seen that the $\hat{S}^2$ operator has one- and two-electron terms. Therefore, the expectation values of it can be expressed through the one- and second-order reduced density matrices. The eigenvalues of the $\hat{S}^2$ operator fulfill the conditions of any angular momentum operator so that

$$\hat{S}^2 f(\omega_1, ..., \omega_N) = S(S+1) f(\omega_1, ..., \omega_N) \tag{1.28}$$

and

$$\hat{S}_z f(\omega_1, ..., \omega_N) = m_S f(\omega_1, ..., \omega_N), \tag{1.29}$$

where S is the total spin quantum number and m_S=*-S,-S+1,...S-1,S*. The $f(\omega_1, ..., \omega_N)$ are the corresponding polielectronic spin eigenfunctions. Their construction for a large number of particles is not a trivial task.[9]

Since the electronic Hamiltonian of the non-relativistic time-independent Schrödinger equation does not depend on the spin coordinates, $\hat{S}^2$ and $\hat{S}_z$ commute with $\hat{H}$. Therefore, the spin quantum numbers S and m_S of a pure spin state can be obtained from the exact eigenvalues of the electronic non-relativistic

time-independent Schrödinger equation $\Psi_i(\vec{x}_1, ..., \vec{x}_N)$. As we have mentioned before, $\hat{H}$ only depends on spatial coordinates, but the wave functions contains spin coordinates to completely describe the electrons. Polyelectronic wave functions are built under the orbital approximation, by which the electrons are described by individual functions (orbitals). The so-called Slater determinant is an antisymmetric wave function of a N-electron system, of the form

$$\Psi(\vec{x}_1, ..., \vec{x}_N) = (N!)^{-\frac{1}{2}} \begin{vmatrix} \varphi_i(\vec{x}_1) & \varphi_j(\vec{x}_1) & \cdots & \varphi_N(\vec{x}_1) \\ \varphi_i(\vec{x}_2) & \varphi_j(\vec{x}_2) & \cdots & \varphi_N(\vec{x}_2) \\ \vdots & \vdots & & \vdots \\ \varphi_i(\vec{x}_N) & \varphi_j(\vec{x}_N) & \cdots & \varphi_N(\vec{x}_N) \end{vmatrix} \tag{1.30}$$

where $\varphi(\vec{x})$ represents the one-electron spinorbitals. In the simplest approach spinorbitals depend on the spatial and spin coordinates as

$$\varphi(\vec{x}) = \varphi(\vec{r})s(\omega), \tag{1.31}$$

where the first term of the r.h.s. is the spatial orbital, and the second is the spin function. Restricted spinorbitals are those with the same spatial part for α and β electrons while in unrestricted spinorbitals the spatial part of electrons α and β can be different. It can be shown that any Slater determinant formed by a set of spinorbitals is always an eigenfunction of the total S_z operator, with eigenvalue m_S=$\frac{1}{2}$(N_α-N_β), where N_α and N_β are the number of α and β electrons, respectively. Any closed-shell restricted Slater determinant, so that N_α=N_β, is also eigenfunction of $\hat{S}^2$ with eigenvalue and $S = 0$. Any high-spin restricted open-shell determinant is also eigenfunction of $\hat{S}^2$ with the eigenvalue S(S+1) and S=M_s=$\frac{1}{2}$(N_α-N_β). Any unrestricted Slater determinant is not in general an eigenfunction of $\hat{S}^2$. It is also important to note that restricted Slater determinants with the same m_S value can be linearly combined to produce spin-adapted configurations that will be eigenfunctions of $\hat{S}^2$.[10]

1.3 Atoms in molecules

Atoms are the building blocks of chemistry but from the perspective of quantum mechanics they are not observables. Since Lewis theory[11] most chemistry has been rationalized using the concept of *an atom in a molecule*. For instance, we know from undergraduate organic courses that carbon binds to itself in many different ways, forming mostly single, double, triple and aromatic bonds. Although the carbon atom is the same building block in all these bonds, the character of each carbon is clearly different. In order to characterize the *electronic structure* of these molecules it is convenient to distinguish and characterize the role of the atoms that form these molecules, *i.e.*, to define *an atom in a molecule*. Therefore, any AIM is a conceptual construct but with an irrefutable utility. By characterizing atoms inside a molecule we are defining an *atomic partition*. An atomic partition (or partitioning) is a well-defined method to subdivide the molecule into its constituting atoms. An atomic partition provides the means to define atomic properties that can be used to (chemically) rationalize the electronic structure of a given molecule. In this sense, atomic domains are a useful tool to define partial charges, partial multipoles and to perform a bonding analysis. Taking advantage of different AIM's definitions, quantum chemists have defined atomic populations, bond orders descriptors,[12–18] energy partitioning,[19,20] aromaticity indexes,[21,22] among other useful tools. The definition of the atom within the molecule is the first step towards an analysis of the wave function in classical chemical terms.

There is not a unique way to define an atomic partition and, to some extent, all the proposals are, in one way or another, arbitrary. Therefore, it is important to know the limitations and the drawbacks of the partition we employ. Over the last decades[23–32] several AIM definitions have been proposed. In this section we will briefly summarize the AIM definitions used in Chapters 3-7. We will put more emphasis on the definition of the so-called Becke atoms because it is the basis for the new AIM proposed in Chapter 3.

1.3.1 Hilbert-space analysis

Within the MO-LCAO (molecular orbitals as linear combination of atomic orbitals) approach, the molecular orbitals (MO) are expanded on a finite set of atomic-centered one-electron functions $\{\chi_\mu(\vec{r})\}$ as

$$\varphi_i(\vec{r}) = \sum_\mu c_{\mu i}\chi_\mu(\vec{r}). \tag{1.32}$$

Where $c_{\mu i}$ is the expansion coefficient of the atomic orbital μ in the molecular orbital i. The set of functions centered on a given atom conform an atomic Hilbert subspace $\{\chi_\mu(\vec{r})\}_{|\mu\in A}$. In this context, an atom can be defined by its nucleus and the subspace of one-electron basis set centered on it. Any molecular quantity can be decomposed in terms of atomic, diatomic or in general poliatomic contributions. For instance, a molecular orbital can be rewritten as a sum of its atomic counterparts as

$$\varphi_i(\vec{r}) = \sum_A \sum_{\mu\in A} c_{\mu i}\chi_\mu(\vec{r}). \tag{1.33}$$

Most interesting is the decomposition of the electronic density. It can be written in terms of the atomic orbital basis as

$$\rho(\vec{r}) = \sum_{\mu\nu} D_{\mu\nu}\chi^*_\nu(\vec{r})\chi_\mu(\vec{r}), \tag{1.34}$$

where $\mathbf{D}$ is the spin-less first-order density matrix in the atomic orbital basis representation. Matrix $\mathbf{D}$ contains contributions from the alpha and beta first-order density matrices, $\mathbf{P}^\alpha$ and $\mathbf{P}^\alpha$, respectively.

$$\mathbf{D} = \mathbf{P}^\alpha + \mathbf{P}^\beta. \tag{1.35}$$

By integrating eq. (1.34) one obtains the total number of electrons of the system

$$N = \int \rho(\vec{r})d\vec{r} = \sum_{\mu\nu} D_{\mu\nu}\int \chi^*_\nu(\vec{r})\chi_\mu(\vec{r})d\vec{r} = \sum_{\mu\nu} D_{\mu\nu}S_{\nu\mu} = \sum_\mu (DS)_{\mu\mu}, \tag{1.36}$$

where matrix **S** is the atomic overlap matrix, with elements

$$S_{\nu\mu} = \int \chi^*_\nu(\vec{r})\chi_\mu(\vec{r})d\vec{r}. \tag{1.37}$$

The summation at the r.h.s. of eq. (1.36) can be grouped according to which atom the basis function is attached to. This naturally permits to express the total number of electrons as a sum of atomic contributions, or Mulliken's[23] atomic gross populations N_A

$$N = \sum_A \sum_{\mu\in A} (DS)_{\mu\mu} = \sum_A N_A. \tag{1.38}$$

Alternatively, one can also write eq. (1.38) in the form

$$N = \sum_A \sum_B \sum_{\mu\in A} \sum_{\nu\in B} D_{\mu\nu}S_{\nu\mu} = \sum_A N_{AA} + \sum_{A\neq B} N_{AB} \tag{1.39}$$

where N_{AA} and $N_{AB} + N_{BA}$ are the Mulliken's[23] net and overlap populations, respectively.

Other quantities can be obtained by integrating two-electron densities. For instance, as proposed by Mayer, by decomposing the exchange density one can get bond orders (BO) and free valences. In Chapter 4 we tackle the decomposition of the spin squared operator under the Hilbert-space framework. The integrations in eq. (1.37) are analytical for Gaussian functions. Therefore, its implementation is straightforward and the computational cost associated with the analysis is very small. Moreover, the errors associated with any numerical integration method are avoided. However, the analysis also presents some drawbacks, namely

- Basis set dependency. It has been observed that the results are moderately depending on the basis set used.[33]

- It shows unphysical results for basis sets without marked atomic character e.g. diffuse basis functions. The use of these functions is mandatory for the correct description of molecular systems, especially those with

partial charges such as anions. Therefore, most Hilbert-space analysis done in molecules suffer from this drawback.[34]

- Ill-defined in the complete basis set limit (CBS).
- Inapplicable beyond the LCAO-MO approximation e.g. if plane waves are used to expand the MO of the system.

However, other approaches can be used to avoid these inconveniences. One alternative is the 3D-space analysis.

1.3.2 Three-dimensional space analysis

In the framework of the 3D-space analysis the atom is identified by a region of the physical space or atomic domain and its nucleus. If the atomic domains are disjoint, then the decomposition of a physical quantity into atomic contributions can be simply carried out by integration over the corresponding atomic domains Ω_A as

$$\begin{aligned} \int f(\vec{r})d\vec{r} &= \sum_A \int_{\Omega_A} f(\vec{r})d\vec{r} \\ \int f(\vec{r}_1,\vec{r}_2)d\vec{r}_1 d\vec{r}_2 &= \sum_{A,B} \int_{\Omega_A}\int_{\Omega_B} f(\vec{r}_1,\vec{r}_2)d\vec{r}_1 d\vec{r}_2, \end{aligned} \tag{1.40}$$

where one can see that one and two-electron quantities quite naturally decompose into one-center and one- and two-center contributions, respectively. Voronoi cells and the topological atom from QTAIM by Bader[26] are examples of 3D-space partitioning with strict boundaries. Bader's QTAIM is based on the topological properties of the electron density.

Topology of the electron density

The density is a continuous function defined at every point of the real (three-dimensional) space. Therefore, it easily renders to a topological analysis. The calculation of the gradient of a function provides the set of critical points (CPs) of the function:

$$\vec{\nabla}\rho(\vec{r}_c) = 0. \tag{1.41}$$

The characterization of these critical points is done through the analysis of the second derivatives of the density at the critical point. All the second derivatives of the density are collected in the so-called Hessian matrix:

$$\mathbf{H}[\rho](\vec{r}_c) = \begin{pmatrix} \frac{\partial^2\rho(\vec{r})}{\partial x^2} & \frac{\partial^2\rho(\vec{r})}{\partial x\partial y} & \frac{\partial^2\rho(\vec{r})}{\partial x\partial z} \\ \frac{\partial^2\rho(\vec{r})}{\partial y\partial x} & \frac{\partial^2\rho(\vec{r})}{\partial y^2} & \frac{\partial^2\rho(\vec{r})}{\partial y\partial z} \\ \frac{\partial^2\rho(\vec{r})}{\partial z\partial x} & \frac{\partial^2\rho(\vec{r})}{\partial y\partial z} & \frac{\partial^2\rho(\vec{r})}{\partial z^2} \end{pmatrix}_{\vec{r}=\vec{r}_c} \tag{1.42}$$

that is a real symmetric matrix and thus can be diagonalized,

$$\mathbf{H}[\rho]\mathbf{L} = \mathbf{L}\mathbf{\Lambda} \tag{1.43}$$

i.e., put in a diagonal form,

$$\mathbf{\Lambda} = \begin{pmatrix} \frac{\partial^2\rho(\vec{r})}{\partial x_1^2} & 0 & 0 \\ 0 & \frac{\partial^2\rho(\vec{r})}{\partial y_1^2} & 0 \\ 0 & 0 & \frac{\partial^2\rho(\vec{r})}{\partial z_1^2} \end{pmatrix}_{\vec{r}_1=\vec{r}_c} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \tag{1.44}$$

where $(\lambda_1 \leq \lambda_2 \leq \lambda_3)$ are the three eigenvalues of the Hessian matrix, *i.e.*, the curvatures. We will label each critical point according to its rank and signature as (ω, σ). The difference between the number of positive and negative eigenvalues is the signature, and the total number of non-zero eigenvalues is the rank. Assuming non-zero eigenvalues we can classify the CP by the sign of its curvatures. Each positive curvature contributes +1 to the signature and every negative curvature adds -1, giving four different CP types:

- **(3,-3)**. **Attractor or Nuclear Critical Point (ACP)**. All the curvatures are negative in a ACP, and thus this CP is a maximum of the electron density. These regions usually coincide with an atomic position and an atom-in-molecule within QTAIM theory is characterized by one and only one ACP. Although it is not usual one may encounter maxima of the electron density which do not coincide with an atomic position; those are known as non-nuclear maxima (NNA).

- **(3,-1)**. **Bond Critical Point (BCP)**. A BCP shows two negative

curvatures and a positive one. The BCP is found between two ACP. The positive eigenvalue (λ_3) corresponds to the direction connecting these two ACP and the negative eigenvalues form a plane in the perpendicular direction. The existence of a BCP is used as an indicator of the presence of a chemical bond between the atoms identified by the two ACPs.

- **(3,+1) Ring Critical Point (RCP)**. A RCP has two positive curvatures and one negative one (λ_1). Its presence indicates a ring structure, which seats in the plane formed by the positive eigenvalues. If the molecule is planar the RCP is located in the minimum of the electron density inside the ring structure.

- **(3,+3) Cage Critical Point (CCP)**. A CCP has three negative eigenvalues and it is thus a minimum of the electron density. Its presence indicates a cage structure and the CCP locates close to its center.

The topology of electron density fulfills the Poincaré-Hopf expression, that gives the relationship that should be fulfilled between the number of critical points:

$$n_{\mathrm{ACP}} - n_{\mathrm{BCP}} + n_{\mathrm{RCP}} - n_{\mathrm{CCP}} = 1. \tag{1.45}$$

In fig. (1.1) we can see the density isocountor in dark blue, and the position of ACPs in red spheres, BCPs in green spheres and the bond path in yellow lines for the water molecule. The gradient lines (perpendicular to the isocountor lines) are depicted in cyan and they all end up in the infinity or in the separatrix surface defining the boundaries of the atomic domains. The *atoms in molecules* as defined by QTAIM have their boundaries limited by the *zero-flux* surface in the gradient vector field of the electron density. This zero-flux surface is given by:

$$\nabla\rho(\vec{r}) \cdot \vec{n}(\vec{r}) = \vec{0} \qquad\qquad \forall \vec{r} \in S(\vec{r}), \tag{1.46}$$

where $n(\vec{r})$ is the unit vector perpendicular to the zero-flux surface $S(\vec{r})$ or separatrix. This surfaces show the boundaries between AIMs which are, therefore, non-overlapping regions.

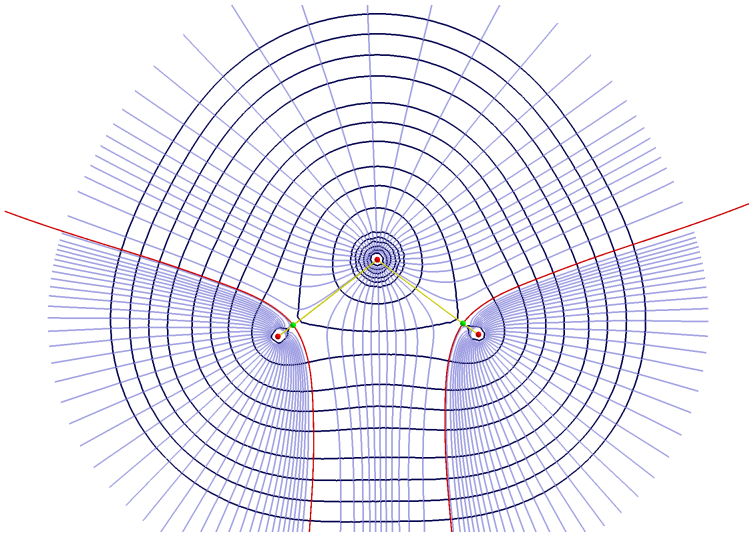

Figure 1.1: Topological analysis of the electron density for the water molecule.

There are two disadvantages to the method, namely the appearance of spurious non-nuclear attractors, which makes the analysis in chemical terms rather difficult, and the computational cost associated to the construction of the complex-shaped atomic domains. As mentioned before, non-nuclear attractors are maxima of the molecular densities not located on any nuclear position.

Fuzzy atoms

An alternative is to consider the so called fuzzy atoms. Within this framework, atoms do not present strict boundaries but they can overlap. Over the last years it has been shown how concepts like bond orders,[27,28,35] overlap populations,[27] valences[27] or energy components,[36] effective atomic orbitals[37] could be generalized to the framework of the "fuzzy" atoms.

In order to treat the different schemes, disjoint atoms or "fuzzy" atoms, in a common framework, one can introduce a non-negative weight function $w_A(\vec{r})$ for each atom and each point of the 3D space satisfying the requirement

$$\sum_A w_A(\vec{r}) = 1. \tag{1.47}$$

In the case of the "fuzzy" atoms, the value of $w_A(\vec{r})$ is large in the vicinity of

the nucleus of atom A and quickly becomes negligible outside. In the special case of disjoint atomic domains, such as those of QTAIM[26] or Voronoi cells, $w_A(\vec{r}) = 1$ for points inside the atomic domain of A and $w_A(\vec{r}) = 0$ outside of it.

The decomposition of a physical quantity into atomic contributions can be performed by inserting the identity (1.47) one or two times for the one and two-electron integrals, respectively

$$\begin{aligned} \int f(\vec{r})d\vec{r} &= \sum_A \int w_A(\vec{r}) f(\vec{r})d\vec{r} \\ \int f(\vec{r}_1,\vec{r}_2)d\vec{r}_1 d\vec{r}_2 &= \sum_{A,B} \iint w_A(\vec{r}_1) w_B(\vec{r}_2) f(\vec{r}_1,\vec{r}_2)d\vec{r}_1 d\vec{r}_2. \end{aligned} \tag{1.48}$$

The first "fuzzy" atom definition was proposed by Hirshfeld.[25] In this scheme the atomic weight of atom A at a given point in the space is determined by the ratio

$$w_A(\vec{r}) = \frac{\rho_A^0(\vec{r})}{\sum_B \rho_B^0(\vec{r})}, \tag{1.49}$$

where $\rho_A^0(\vec{r})$ represents the promolecular density of the atom A.

In the classical Hirshfeld definition the resulting shape of the atoms in the molecule are strongly dependent on the choice of the promolecular state of the atom. Bultinck *et al.*[29] proposed an improved Hirshfeld-Iterative scheme that corrects this dependency by an iterative process to obtain promolecular atomic densities that integrate to the same (usually fractional) number of electrons as do the atoms in the molecule

$$\int \rho_A^0(\vec{r})d\vec{r} = N_A \neq Z_A, \tag{1.50}$$

where Z_A is the atomic number of atom A and N_A is the actual population of the atom in the molecule.

An alternative to compute the atomic weight was proposed by Salvador and Mayer[27] making use of the Becke atoms; *i.e.,* the fuzzy atomic Voronoi cells introduced by Becke in the context of his celebrated multicenter numerical integration scheme.[38]

Becke's atomic weights, $w_A(\vec{r})$, are algebraic functions which strictly satisfy the sum rule (1.47). In addition, they also fulfill $w_A(\vec{R}_A) = 1$ and $w_B(\vec{R}_A) = 0$. That is, the region (point) of the space occupied by the nucleus is fully associated to its corresponding atom.[a]

The scheme was formulated as follows. Let us consider two nuclei, A and B, at a distance R_{AB}. For any point in the space, $\vec{r}$, one can calculate the following quantity

$$\mu_{AB} = \frac{r_A - r_B}{R_{AB}}, \tag{1.51}$$

where r_A and r_B represent the distances of that point to nucleus A and B, respectively. The values of μ_{AB} for different points of the space are summarized in fig. (1.2). The surface $\mu_{AB} = 0$ corresponds to the perpendicular plane that bisects the internuclear axis (*i.e.,* the face of the Voronoi cell), whereas $\mu_{AB} = -1$ and $\mu_{AB} = 1$ conform with the line extending from nucleus A to infinity and from nucleus B to infinity, respectively. Thus, the points of the space for which $\mu_{AB} < 0$ are within the Voronoi cell of atom A. That of atom B is formed by those where $\mu_{AB} > 0$. The following step function

$$s_A(\mu_{AB}) = \begin{cases} 1 & -1 \leq \mu_{AB} \leq 0 \\ 0 & 0 < \mu_{AB} \leq 1 \end{cases} \tag{1.52}$$

can be used to define the Voronoi cell of atom A in this case.

In order to generalize the scheme for a general system formed by a set of atoms one needs to compute the μ_{AB} and step functions for all pair of atoms, and

[a] This is not necessarily the case for Hirshfeld based fuzzy atom definitions.

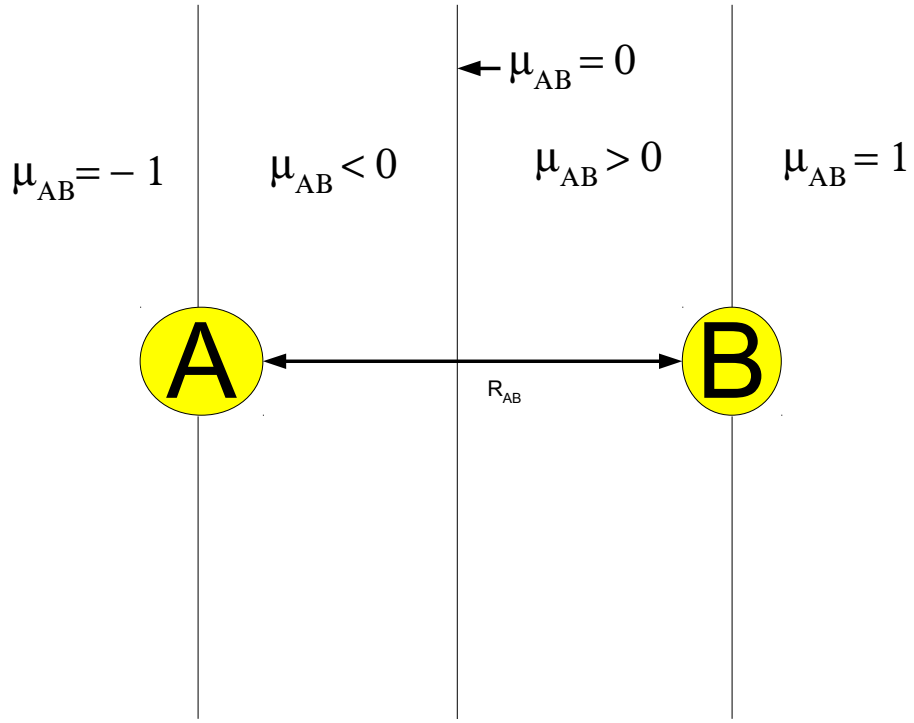

Figure 1.2: Schematic representation of the confocal elliptical coordinate μ_{AB}. for atoms A and B.

construct a weight function to define the atomic Voronoi cell of atom A as

$$w_A(\vec{r}) = \frac{P_A(\vec{r})}{\sum_B P_B(\vec{r})} \tag{1.53}$$

where

$$P_A(\vec{r}) = \prod_{B \neq A} s_A(\mu_{AB}). \tag{1.54}$$

The step function (1.52) can be replaced by a continuous, monotonically decreasing function in the range (-1,1), fulfilling the requirements $s(-1) = 1$ and $s(+1) = 0$ in order to define the fuzzy Voronoi cells. For that purpose Becke suggested the simple polynomial function

$$s_A^k(\mu_{AB}) = \frac{1}{2}[1 - f_k(\mu_{AB})] \tag{1.55}$$

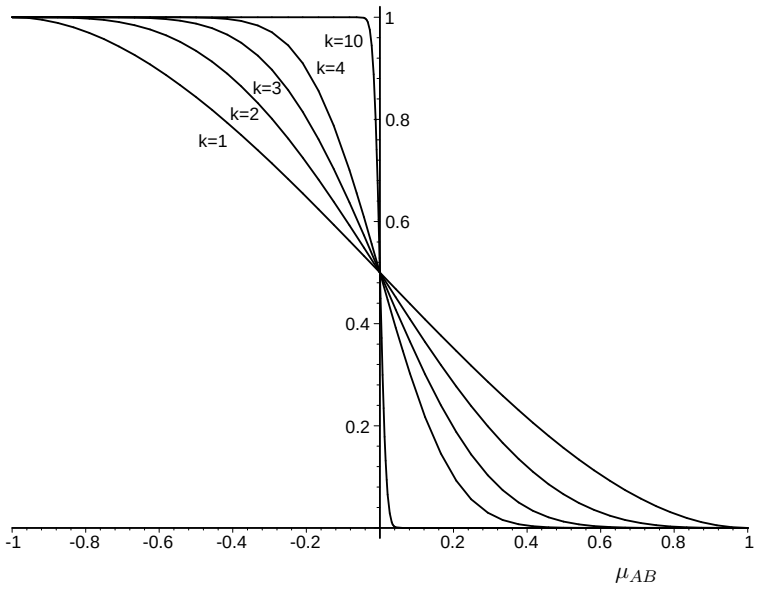

Figure 1.3: Becke's cutoff profiles of the weight function for different values of the stiffness parameter k. Atom A to the left, B to the right.

where

$$f_1(\mu) = \frac{3}{2}\mu - \frac{1}{2}\mu^3 \tag{1.56}$$

and

$$f_k(\mu) = f[\, f_{k-1}(\mu)] \tag{1.57}$$

The integer k is known as the *stiffness* parameter and controls the decay (see fig. (1.3)) of the atomic weight from the value of one at the nuclear position to zero at the position of the neighboring atoms. Mathematically speaking, k determines the order of the polynomial function of eq. (1.57).

Figure (1.3) depicts such profile as a function of k. The larger the value of k the steeper is the cutoff profile. For larger values of k the shape of the step-function (eq. (1.52)) is reproduced.

The fuzzy Voronoi cells thus defined do not account for the different atomic sizes in heteronuclear systems, as the faces of the Voronoi cells exactly bisect the internuclear axis between the two neighboring atoms. In order to define a shifted cutoff profile one can use the same cutoff function (1.55) but using as argument the following transformed coordinate

$$\nu_{AB} = \mu_{AB} + a_{AB}(1 - \mu_{AB}^2), \tag{1.58}$$

where $-1/2 \leq a_{AB} \leq 1/2$ in order to ensure that $-1 \leq \nu_{AB} \leq 1$. The position of the cell boundary is now controlled by the value of the parameter a_{AB}, which in the new coordinates is given by the condition $\nu_{AB} = 0$.

At the point where the shifted cell boundary intersects the interatomic plane one has

$$\mu_{AB} = \frac{r_A - r_B}{r_A + r_B} \tag{1.59}$$

Becke suggested to relate the ratio r_A/r_B with the relative size of the atoms A and B making use of some set of reference *fixed* atomic radii R_A^0 and R_B^0 as

$$\frac{r_A}{r_B} = \frac{R_A^0}{R_B^0} = \chi_{AB} \tag{1.60}$$

Substituting (1.60) into (1.59) and using the condition $\nu_{AB} = 0$ one gets the following relationship between a_{AB} and χ_{AB}

$$a_{AB} = \frac{1 - \chi_{AB}^2}{4\chi_{AB}}. \tag{1.61}$$

In fig. (1.4) one can see how the cell boundary is shifted from the midpoint of the internuclear axis when the radius of the atom on the left is twice as large ($\chi = 2$) as that of the atom on the right.

To summarize, the shape of the Becke atoms (the value of the atomic weights) is fully determined by the set of fixed atomic radii used and the stiffness

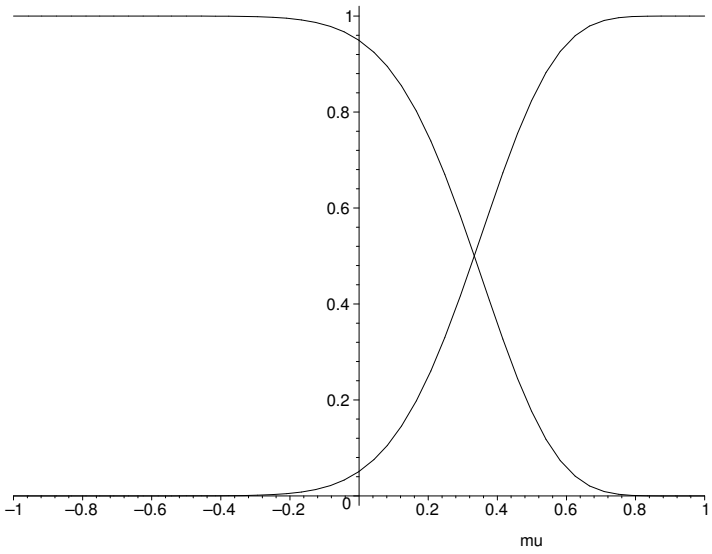

Figure 1.4: Shifted interatomic profiles of the weight function for $k = 3$ and $\chi_{AB} = \frac{R_A}{R_B} = 2$. Atom A to the left, B to the right.

parameter. The former determines the size of the atomic Voronoi cells, whereas the latter controls the shape of the cutoff profile, *i.e.,* to which extent two neighboring atoms share the physical space between them. In the original paper, Becke used the set of empirical atomic radii of Bragg and Slater[39] and $k = 3$, on the basis of a better performance of his numerical integration scheme.

An alternative to the use of the fixed set of empirical radii is to use an internal criterion to determine the position of the cell boundaries between all neighboring atoms. Thus, instead of a set of fixed radii one needs to establish a strategy to define the set of χ_{AB} values, where A and B represent any pair of atoms sharing a cell boundary. Mayer *et al.*[27] proposed a new criterion based on the position of the extremum (typically a minimum) of the density along the internuclear axis connecting every two neighboring atoms to determine the location of the corresponding interatomic cell boundary. That scheme was later referred to as Becke-ρ,[28] and can be considered as a good and efficient adaptation of Bader's partitioning.

1.3.3 Effective atomic densities

The formalism of the atomic and diatomic effective densities was first introduced by Vyboishchikov *et al.*[40] The main feature of the formalism is that it allows, in a common framework, the derivation of the working expressions for atomic populations, bond orders, atomic valences, molecular energy components, *etc.* for any kind of AIM definition. Within this formalism, the one- and two-electron densities are *exactly* decomposed into components that can be considered their one-center and one- and two-center contributions, respectively.

In the simplest case, one can define the effective atomic contributions to the electron density, $\rho_A(\vec{r})$, simply fulfilling $\rho(\vec{r}) \equiv \sum_A^N \rho_A(\vec{r})$. The integral over the whole space of this function for atom A quite naturally yields the electron population associated to the atom

$$\int \rho_A(\vec{r}) d\vec{r} = N_A. \tag{1.62}$$

The actual numerical value depends upon how $\rho_A(\vec{r})$ is defined. In the framework of 3D-space analysis, $\rho_A(\vec{r})$ can be written in general as

$$\rho_A(\vec{r}) = w_A(\vec{r})\rho(\vec{r}), \tag{1.63}$$

where w_A is the non-negative weight function fulfilling conditions given in eq. (1.47). The actual definition of *atom in the molecule* ("fuzzy" or disjoint) is contained in the shape of the atomic weight functions.

In the case of Hilbert-space analysis, the effective atomic density can be most suitable written in terms of the matrix elements of the LCAO density matrix as

$$\rho_A(\vec{r}) = \sum_{\mu \in A} \sum_{\nu} D_{\mu\nu} \chi_\nu^*(\vec{r}) \chi_\mu(\vec{r}). \tag{1.64}$$

It is trivial to see that the integration of eq.(1.64) yields Mulliken's gross population of atom A as defined in eq. (1.38).

One advantage of using the formalism of effective atomic densities is that one can switch from 3D-space to Hilbert-space formulae or *vice versa* simply by taking the appropriate form of the effective densities involved in the calculation.

1.3.4 Mapping between 3D and Hilbert-space analyses

A correspondence between 3D-space and Hilbert-space analysis can also be established by introducing a mapping between the atomic overlap matrices (used in the framework of 3D-space analysis) and their Hilbert-space counterparts. [18] Let us consider for simplicity a closed-shell system with doubly occupied molecular orbitals. In the framework of 3D-space analysis, the gross atomic population of atom A is obtained as

$$N_A = \int w_A(\vec{r})\rho(\vec{r})d\vec{r} = \sum_{\mu\nu} D_{\mu\nu} S^A_{\nu\mu} = 2\sum_{\mu\nu}\sum_{i}^{occ} c^*_{\nu i} S^A_{\nu\mu} c_{\mu i} = 2\, tr(\mathbf{C}^+\mathbf{S}^{\mathbf{A}}\mathbf{C}), \tag{1.65}$$

where

$$S^A_{\nu\mu} = \int w_A(\vec{r})\chi^*_\nu(\vec{r})\chi_\mu(\vec{r})d\vec{r} \tag{1.66}$$

are the elements of the atomic overlap matrix $\mathbf{S}^A$ and $\mathbf{C}$ is the matrix containing the orbital coefficients of the occupied molecular orbitals. The Mulliken gross population can be written in terms of the molecular orbital coefficients as

$$N_A = 2\sum_{\mu\in A}\sum_{\nu}\sum_{i}^{occ} c^*_{\nu i} S_{\nu\mu} c_{\mu i} = 2\, tr(\mathbf{C}^+\mathbf{S}\eta^{\mathbf{A}}\mathbf{C}) \tag{1.67}$$

where $\eta^{\mathbf{A}}$ is a block-truncated unit matrix with all elements equal to zero except $\eta^A_{\mu\mu} = 1$ for $\mu \in A$. Comparing eqs. (1.65) and (1.67) one can put into correspondence with the atomic overlap matrix in atomic orbital basis $\mathbf{S}^A$ the matrix product $\mathbf{S}\eta^A$. Note that this Mulliken's *effective* atomic overlap matrix is non symmetric.

When looking for an AIM, we are not only simply interested in a subdivision of the space, but rather in splitting global properties into the individual atoms

and atoms pairs. Taking advantage of the AIM definitions we can describe for instance local spins within a molecule. In the next section the origins of the local spin analysis will be reviewed.

1.4 Local spin

In the last years there has been a growing interest in recovering local spins from ab initio wave functions.[41–52] Clark and Davidson proposed the first definition of the local spins in 2001.[41] They were motivated by the phenomenological Heisenberg Hamiltonian

$$\hat{H} = -\sum_{A<B} J_{AB}\hat{S_A}\hat{S_B}, \tag{1.68}$$

where J_{AB} is the so-called exchange coupling constant and $\hat{S_A}$ and $\hat{S_B}$ are atomic spin operators. Usually, when the Heisenberg Hamiltonian[53,54] is applied, the value of the local spins and the atoms to which the spins are located are assigned in a knowledge-based manner. They proposed an alternative approach by computing quantities like $< \hat{S_A}\hat{S_B} >$ from *ab initio* wave functions. They defined local spin operators $\hat{S}_A$ by projecting the overall spin operator $\hat{S}$. These local operators are consistent with the definition of angular momentum operators. The expectation values of $\hat{S}_A$ and $\hat{S_A}\hat{S_B}$, are considered as the atomic and diatomic contributions to the overall $\langle \hat{S}^2 \rangle$ value, respectively. The formalism was applied to diatomic molecules,[41,42] hydrocarbons,[41] transition metal complexes,[41,42] and organic diradicals.[41–43]

This formalism leads to $< \hat{S_A}\hat{S_B} >= -\frac{3}{8}BO_{AB}$ (where BO stands for the the Mayer bond order) and $< \hat{S_A} >= \frac{3}{8}\sum_{B\neq A} BO_{AB}$ for single determinant closed-shell systems. That is, the diatomic term is proportional to the BO of the given pair of atoms, and the atomic contribution is proportional to the sum of the BO involving the given atom. Even for genuine diamagnetic systems treated at the single determinant level the formalism attributes some local spins.

This unphysical result provided by the Clark and Davidson approach moti-

vated Mayer to propose an alternative.[44] The new formalism was based on the partition of the expectation value of the $\hat{S}^2$ operator in the spirit of classical population analysis. The expectation value of the $\hat{S}^2$ operator is decomposed as a sum of atomic and diatomic contributions:

$$\langle \hat{S}^2 \rangle = \sum_A \langle \hat{S}^2 \rangle_A + \sum_{A,B \neq A} \langle \hat{S}^2 \rangle_{AB}. \tag{1.69}$$

In the case of single-determinant wave functions Mayer obtained expressions for the atomic and diatomic contributions to $\langle \hat{S}^2 \rangle$ depending only on the molecular spin-density matrix. Such a decomposition trivially leads to zero atomic spins for any closed-shell restricted wave function. This decomposition was carried out originally in the framework of the Hilbert-space analysis. The generalization of the formulae to the 3D physical-space analysis was already outlined by Mayer[44] and was explored by Alcoba et al.[48] using both Bader's QTAIM[26] domains and Becke's "fuzzy" atoms.[38]

The first attempt for a $\langle \hat{S}^2 \rangle$ decomposition scheme (eq. (1.69)) for a general wave function , *i.e.* beyond a single determinant description, was proposed by Alcoba *et al.*[45] The components of $\langle \hat{S}^2 \rangle$ were expressed in terms of the total spin-density matrix and therefore zero spins were obtained for every singlet state system, even for singlet dirradicals. Therefore, using this formalism, a genuine diamagnetic molecule could not be distinguished from a singlet diradical or an antiferromagnet. For such cases, if treated with a proper wave function, one may assume the existence of local spins even though the global system is a singlet.

Later on, Mayer[46] proposed an alternative formulation for a general wave function on the basis of several physical requirements, namely:

i) to obtain zero local spins for closed-shell restricted wave functions.

ii) proper asymptotics, *i.e.*, in the dissociation limit the atoms/fragments should have the same local $\langle \hat{S}^2 \rangle$ value as the respective free atoms/fragment radical moieties.

iii) the formula used for a general wave function should reduce to that used in the single-determinant case if applied to single-determinant wave functions.

In his formulation, the atomic and diatomic components of $\langle \hat{S}^2 \rangle$ were expressed in terms of the value that they would have in the single-determinant case, plus correction terms arising form the deviation of the first-order density matrix from the indempotency, as well as contributions from the cumulant of the second-order density matrix (both vanish for single determinant wave functions thus fulfilling requirement *iii*.). These correction terms were distributed between atomic and diatomic contributions in such a way fulfillment of condition *ii* was guaranteed. The Hilbert-space realization of the formalism given in ref. 47 corroborated the physical expectations in all cases.
For instance, small atomic contributions were obtained for every closed-shell molecule at equilibrium distance, including *e.g.* the carbon atoms of singlet π-conjugated systems such as benzene calculated at the CASSCF level of theory. At the same time, large local spins were found for the square cyclobutadiene which is known to be a "molecular antiferromagnet". Furthermore, the CASSCF dissociation curves of O_2 molecule for both singlet and triplet states lead to atomic local-spin values that tend asymptotically to 2, as expected for the $^3P_2$ state of the free oxygen atom. Similarly, ethylene dissociates into two triplet methylene radicals, and so forth.

Although it seemed as there was no further freedom to choose another local-spin decomposition scheme, Alcoba et al.[49] showed that this was not the case. These authors introduced yet another requirement for the partitioning of $\langle \hat{S}^2 \rangle$ for non-singlet states, namely

iv) the one and two-center components should be independent of the S_z value (in the absence of a magnetic field).

Their local spin formulation scheme used spin-free quantities such as the density of effectively unpaired electrons[55] and the "spin-free cumulant matrix",

which depend on the spin-free first- and second-order reduced density matrices.[56,57] Their Hilbert-space local-spin results are similar to, but not identical with those obtained by Mayer and Matito,[47] and keep the physical requirement of S_z independence. But most importantly, the Mayer and Acoba formulations gave different results **even for singlet states**, where the dependence on the spin density of the Mayer formulation vanished. Therefore, it was clear that Alcoba's formulation was not merely correcting for the dependence of the Mayer expressions on the spin density, but it corresponded actually to a different approach. This motivated us to deeply investigate both formulations in order to find any connection between them. In Chapters 4 these approaches are presented and discussed.

1.5 Effective atomic orbitals

The effective atomic orbitals (eff-AOs)[4,37,58–61] were introduced by Prof. Mayer with the aim of recovering the classical concept of electron configuration (1s, 2s, 2p orbitals or their hybrids) of an atom within a molecule from the *a posteriori* analysis of the wave function. The formalism was first proposed almost two decades ago in the framework of Hilbert-space analysis,[58] and then it was generalized for an arbitrary Hermitian bilinear "localization functional".[59] More recently it was applied to the case of "fuzzy" atoms.[37] It had also been formulated in the context of Bader's QTAIM,[60] but that version was never actually realized. We will tackle this problem in Chapter 1.5.

By performing the eff-AOs analysis one obtains for each atom of the molecule a set of orthogonal[b] hybrids and their respective occupations. What makes this procedure appealing is the fact that independently of the basis set used in the *ab-intio* calculation, one obtains as much populated atomic hybrids as orbitals in a minimal basis set[c]. Therefore, an "effective minimal basis set" can be extracted from a wave function, even if plane-waves are used to expand

[b]The hybrids belonging to different atoms are not in general orthogonal to each other except if the atoms are defined by strict boundaries.

[c]Special provisions should be made in the case of hypervalent systems

the wave function of the system.[61]

In order to shortly introduce the formulation of the eff-AOs let us consider a system with n orthonormalized doubly occupied orbitals $\varphi_i(\vec{r}),\ \ i = 1, 2 \ldots, n$. For each atom A of the molecule one can define the intra-atomic part of every MO within the general 3D-space framework as

$$\varphi_i^A(\vec{r}) \equiv \varphi_i(\vec{r}) w_A(\vec{r}) \tag{1.70}$$

where $w_A(\vec{r})$ is a non-negative weight function which defines the fuzzy domain of atom A in the molecule. One can build a $n \times n$ overlap matrix of the "intraatomic" MOs, $\mathbf{Q}^A$, with elements

$$Q_{ij}^A = \int \varphi_i^{A^*}(\vec{r}) \varphi_j^A(\vec{r}) d\vec{r}. \tag{1.71}$$

The Hermitian matrix $\mathbf{Q}^A$ is diagonalized by the unitary matrix $\mathbf{U}^A$

$$\mathbf{U}^{A,\dagger} \mathbf{Q}^A \mathbf{U}^A = diag\{\lambda_i^A\}. \tag{1.72}$$

The eff-AOs for atom A are obtained as a linear combination of the intraatomic part of the MOs as

$$\chi_i^A(\vec{r}) = \frac{1}{\sqrt{\lambda_i^A}} \sum_{\mu}^{n} U_{\mu i}^A \varphi_\mu^A(\vec{r}) \qquad i = 1, n^A \tag{1.73}$$

where n^A is the number of non-zero eigenvalues λ_i^A. The latter are the corresponding occupation numbers of the eff-AOs, with $0 < \lambda_i^{A,\sigma} \leq 1$.
The eff-AOs of atom A can also be obtained from the diagonalization of the matrix $\mathbf{PS}^A$, where $\mathbf{P}$ is the LCAO density matrix and $\mathbf{S}^A$ is the atomic overlap matrix in the actual (AO or MO) basis.[59] This alternative also permits the straightforward generalization of the method both to the unrestricted single-determinant case and the correlated level.

In the case of non-overlapping or disjoint atoms and for single-determinant wave functions, the localized orbitals obtained are identical to those of the domain averaged Fermi holes (DAFH) analysis before the isopycnic transforma-

tion. In the eff-AOs analysis, however, the relevant objects are the truncated localized orbitals.

The concept of DAFH was introduced by Ponec some years ago in order to shed light on the nature of chemical bonding.[62–75] The original definition of the DAFH was formulated in terms of the conditional probability and the concept of Fermi holes but it can also be derived from the exchange-correlation density. For each atom or fragment in the molecule one can define the DAFH density as

$$g(\vec{r}) = \int_{\Omega_A} \rho(\vec{r}_1, \vec{r}_2) d\vec{r}_2 - \int_{\Omega_A} \rho(\vec{r}_1)\rho(\vec{r}_2) d\vec{r}_2. \tag{1.74}$$

The first step after having determined the DAFH functions consists in the diagonalization of the matrix that represents $g(\vec{r})$ in a proper basis. In the second step, the eigenvalues and eigenvectors resulting from the previous diagonalization are subjected to a isopycnic[76] localization.

1.6 Oxidation states from first principles

The concept of oxidation state (OS) is widespread in transition metal chemistry and in the study of redox and catalytic reactions. The reactivity, spin-state, spectroscopic and geometrical features of transition metal complexes are often rationalized on the basis of the oxidation state of the metal center. According to the IUPAC,[77] the formal OS is typically defined as "the charge it would bear if all the ligands were removed along with the electron pairs that were shared with the central atom". Jørgensen[78] proposed that an OS derived by a d^n configuration, in principal a measurable quantity (by Mossbauer or Raman spectroscopy), should be called physical or spectroscopic OS. In some cases formal and spectroscopic OS coincide but in complicated bonding situations involving non-innocent ligands or in intermediates or transition states of catalytic reactions the formal and the physical OS may diverge. For instance in fig. (1.5) the two situations are depicted for a phenolato/phenoxyl iron complex.[79]

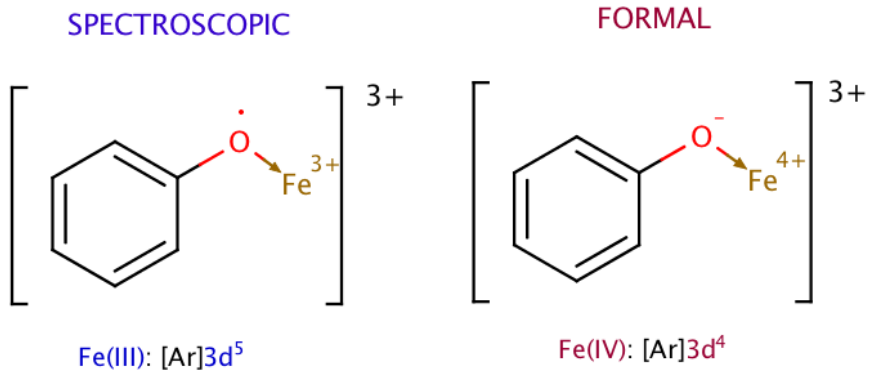

Figure 1.5: Schematic representation of an Fe atom with a non-innocent ligand.

Geometrical parameters have been used to derive OS from empirical data, for instance by using the Bond-Valence-Sum method.[80–82] This method uses empirically determined metal-ligand distances of complexes with well-known OS to extrapolate the OS of other systems by their metal-ligand distances.

The oxidation state is intrinsically related to electronic distribution. However, as for the case of AIM or local spin, OS is not an observable according to quantum mechanics. Therefore, there are no rigorous definitions of them from electronic structure calculations. Several methods to determine OS from first principles have appeared in the literature.[83–87] For instance, populations analysis have been widely used,[3,88] but even though electronic populations do change with oxidation/reduction of the metal center, they are only a pointer of the oxidation state. One can also use spin populations but they need a previous knowledge of the electronic structure (spin state), and, in addition, they are clearly futile in pure singlet states.

The Mulliken and Becke-ρ charges of a series of octahedral complexes (see table (1.1)) are gathered in table (1.2). Only for some of the systems (marked in green) Becke-ρ charges approximate the well-known OS of these systems.

Thom et al.[83] suggested a more sophisticated method based on the decomposition of localized molecular orbitals into atomic contributions, the so-called localized orbital bonding analysis (LOBA) approach. They used different orbital localization methods such as, Boys, Pipek-Mezey, and Edmiston-Ruedenberg

Table 1.1: Set of octahedral complexes and their oxidation states

V^{II}	$[VCl_6]^{4-}$	$[V(H_2O)_6]^{2+}$	$[V(CN)_6]^{4-}$	$[V(CO)_6]^{2+}$
Mn^{II}	$[MnCl_6]^{4-}$	$[Mn(H_2O)_6]^{2+}$	$[Mn(CN)_6]^{4-}$	$[Mn(CO)_6]^{2+}$
Mn^{III}	$[MnCl_6]^{3-}$	$[Mn(H_2O)_6]^{3+}$	$[Mn(CN)_6]^{3-}$	$[Mn(CO)_6]^{3+}$
Fe^{II}	$[FeCl_6]^{4-}$	$[Fe(H_2O)_6]^{2+}$	$[Fe(CN)_6]^{4-}$	$[Fe(CO)_6]^{2+}$
Fe^{III}	$[FeCl_6]^{3-}$	$[Fe(H_2O)_6]^{3+}$	$[Fe(CN)_6]^{3-}$	$[Fe(CO)_6]^{3+}$
Ni^{II}	$[NiCl_6]^{4-}$	$[Ni(H_2O)_6]^{2+}$	$[Ni(CN)_6]^{4-}$	$[Ni(CO)_6]^{2+}$
Zn^{II}	$[ZnCl_6]^{4-}$	$[Zn(H_2O)_6]^{2+}$	$[Zn(CN)_6]^{4-}$	$[Zn(CO)_6]^{2+}$

orbitals. A population analysis of the resulting localized orbitals is then used to assign electrons to the different atoms. In order to study complicated situations the authors performed a calibration to define a threshold to decide whether or not a given electron is localized on a metal center.

Table 1.2: Mulliken (left) and Becke-ρ (right) charges for a set of transition metal complexes.

	Cl^-	H_2O^{HS}	H_2O^{LS}	CN^-	CO	Cl^-	H_2O^{HS}	H_2O^{LS}	CN^-	CO
V^{II}	0.98	1.12	-	0.05	0.64	1.60	1.76	-	1.60	1.64
Mn^{II}	1.10	1.24	1.18	0.10	0.64	1.36	1.64	1.70	1.53	1.53
Mn^{III}	0.93	1.58	1.52	0.35	0.80	1.46	2.05	2.09	1.58	1.67
Fe^{II}	0.86	1.22	1.15	0.01	0.51	1.27	1.63	1.77	1.46	1.44
Fe^{III}	0.99	1.64	1.48	0.24	0.66	1.44	2.06	1.98	1.49	1.58
Ni^{II}	0.99	-	1.08	-0.19	0.31	1.27	-	1.59	1.24	1.30
Zn^{II}	1.02	-	1.06	-0.03	0.52	1.25	-	1.45	1.15	1.19

A few years later, Sit et al.[85] presented a new method based on the separation of the actual occupation of the metal d-type orbitals from the condensed metal charge. A 5x5 matrix for each spin is constructed by projecting the wave functions onto the d-type AOs of the central metal atom. This matrix is then diagonalized and its eigenvalues indicate the number of occupied d-type orbitals assigned to the metal. The procedure has been applied to a series of TM complexes but it can not be employed to systems with metal-metal bonds.

Doubtless there is need for more computational approaches to deal with difficult cases. In this thesis we propose a new formalism based on the definition of the physical OS. That is explored in Chapter 6 by taking advantage of the eff-AOs.

Chapter 2

Objectives

The main objective of this thesis is the development and application of new tools for the analysis of the electron distribution in molecules, focusing on the concepts of local spins, and oxidation state. The precedents of the local spin analysis and the derivation of oxidation states from *ab initio* calculations have been introduced in Chapter 1. In this chapter the detailed objectives of the thesis will be presented.

The concept of the atom within the molecule is crucial to recover chemical concepts from first principles. One of the most accepted in the literature are that provided by the quantum theory of atoms in molecules (QTAIM), by Bader. The relatively high computational cost associated with the integration over the atomic domains in large systems represents a limitation of this atomic definition. Moreover the presence of spurious non-nuclear attractors in some cases restricts its use. **The first aim of this thesis is to propose a new atom in molecule definition based on the "fuzzy" atoms being able to perform similar to the QTAIM topological atom**.

Several definitions of the concept of local spin have appeared in the literature in the last years. As there are in principle infinite ways to decompose $\langle \hat{S}^2 \rangle$ into atomic and diatomic contributions a series of physical requirement have been imposed to restrict the decomposition. **The second goal of this thesis is to establish a relationship among the different definitions of the local spin proposed to date and to introduce new physical requirement**

in order to find a unique, physically sound definition of the local spin.

The formalism of the local spin can be realized in both the Hilbert-space and the three-dimensional real space. There is an apparent ambiguity in decomposing two-electron quantities in the framework of Hilbert-space analysis, in particular for the Mulliken-type scheme. **The third objective of the thesis is to generalize the formalism of the decomposition of $\langle \hat{S}^2 \rangle$ in the framework of three-dimensional space and Hilbert-space analysis, clarifying the ambiguity that appears in the latter**.

The local spin analysis is a relatively novel and unemployed descriptor of the electron distribution in molecules. In order to gain a deeper understanding of the information that the local spin analysis can provide, **the fourth objective of the thesis is to explore the relation between the local spin and the chemical bond.**

The presence of local spins can be intuitively related to the presence of radical centers in a molecule. **The fifth objective of the thesis is to explore the application of the local spin analysis to detect and quantify polyradical character.**

The last part of the thesis is devoted to the derivation of oxidation states from the analysis of the wave function. The concept of eff-AOs was introduced with the aim of recovering the atomic electron configurations from an *a posteriori* analysis of an *ab initio* calculation. The formalism of the eff-AOs within QTAIM was proposed some years ago by Mayer but until now it was never numerically realized. **The sixth goal of this thesis is to compute for the first time the eff-AOs within QTAIM and to describe the prominent role of these orbitals**. The classical chemical concept of oxidation states can be naturally obtained through the electron configuration of the atom within the molecule, **the seventh goal of the thesis is to define a simple and general strategy to obtain oxidation states by using the information provided by the eff-AOs**.

Chapter 3

On the definition of a new "fuzzy" atom scheme

3.1 An approximation to Bader's topological atom

THE JOURNAL OF CHEMICAL PHYSICS **139**, 071103 (2013)

Communication: An approximation to Bader's topological atom

Pedro Salvador and Eloy Ramos-Cordoba[a)]
Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, 17071 Girona, Spain

(Received 16 July 2013; accepted 5 August 2013; published online 19 August 2013)

A new, more flexible definition of fuzzy Voronoi cells is proposed as a computationally efficient alternative to Bader's Quantum Theory of Atoms in Molecules (QTAIM) partitioning of the physical space for large-scale routine calculations. The new fuzzy scheme provides atomic charges, delocalization indices, and molecular energy components very close to those obtained using QTAIM. The method is flexible enough to either ignore the presence of spurious non-nuclear attractors or to readily incorporate them by introducing additional fuzzy Voronoi cells. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4818751]

The identification of an atom within the molecule is a crucial concept that permits the decomposition of physical quantities into atomic (and often also diatomic) contributions. Bader's Quantum Theory of Atoms in Molecules (QTAIM)[1] represents nowadays one of the most widely used definitions of atom in the molecule. In QTAIM, the topology of the electron density is cleverly used to derive the atomic basins and the sharp boundaries that define them. The resulting atoms in the molecule possess some particular properties, namely, the virial theorem is fulfilled within each basin of the system. Such topological atoms are extensively used in the literature for different applications, from a number of different electron population/distribution[2] analysis to molecular energy components[3,4] among others.

There are two disadvantages of the method. First, the analysis of the electron density may lead to so called non-nuclear attractors (NNA), i.e., basins that have no nucleus associated to it. Whereas such objects do exist in condensed phase[5,6] or molecules such as Li_2,[7] a particular combination of level of theory and basis set may give rise to spurious NNAs in dipole-bound water cluster anions[8] or simple molecules such as C_2H_2[3] or Si_2H_2.[9] The existence of such spurious NNAs makes any analysis in chemical terms rather difficult and ambiguous.[9,10] Second, the detailed exploration of the topology of the electron density and specially the cumbersome integration over the complex-shaped atomic domains may be a rather time consuming procedure, as compared with alternative definitions of atom in the molecule. The systematic use of the method in quantum chemical applications of systems of growing complexity such as fullerenes may be compromised.

An alternative to QTAIM within 3D-space analysis are atomic definitions that introduce overlapping or "fuzzy" boundaries. Such "fuzzy" atoms are best represented by introducing a non-negative weight function $w_A(\vec{r})$ for each atom A and each point of the 3D space, which satisfies the requirement

$$\sum_A w_A(\vec{r}) = 1 \tag{1}$$

everywhere. It is assumed that the atomic weight function is large "inside" of atom A and small "outside." Of course, in the special case of QTAIM, $w_A(\vec{r}) = 1$ for points inside the atomic basin of atom A and $w_A(\vec{r}) = 0$ outside of it.

Over the last years we have shown that many quantities such as bond orders, overlap populations,[11] energy components,[12] effective atomic orbitals, or local spins could be generalized and computed in the framework of the "fuzzy atoms." For that mere purpose we have often made use of the simplest Becke's atoms; i.e., the fuzzy atomic Voronoi cells introduced by Becke[13] in the context of his celebrated multicenter numerical integration scheme.

Becke's atoms were originally devised for effective numerical integration of three-dimensional functions of marked atomic character. Becke's atomic weights, $w_A(\vec{r})$, are algebraic functions, which strictly satisfy the sum rule of Eq. (1). In addition, they also fulfill $w_A(\vec{R}_A) = 1$ and $w_B(\vec{R}_A) = 0$ for $B \neq A$; that is, the region (point) of the space occupied by the nucleus is fully associated to its corresponding atom.

The scheme can be formulated as follows. Let us consider two nuclei, A on the left-hand side and B on the right-hand side, at a distance R_{AB}. For any point of the space, $\vec{r}$, one can calculate the following quantity

$$\mu_{AB} = \frac{r_A - r_B}{R_{AB}}, \tag{2}$$

where r_A and r_B represent the distance of that point to nucleus A and B. The surface $\mu_{AB} = 0$ corresponds to the perpendicular plane that bisects the internuclear axis (i.e., the face of the Voronoi cell), whereas $\mu_{AB} = -1$ and $\mu_{AB} = 1$ conform with the line extending from nucleus A to infinity and from nucleus B to infinity, respectively. Thus, the points of the space for which $\mu_{AB} < 0$ are within the Voronoi cell of atom A. The

a)pedro.salvador@udg.edu

following step function

$$s_A(\mu_{AB}) = \begin{cases} 1 & -1 \le \mu_{AB} \le 0 \\ 0 & 0 < \mu_{AB} \le 1 \end{cases} \tag{3}$$

can be used to define the Voronoi cell of atom A.

In order to generalize the scheme for a system formed by a set of atoms, one can define the atomic Voronoi cell of atom A as

$$w_A(\vec{r}) = \frac{P_A(\vec{r})}{\sum_B P_B(\vec{r})}, \quad \text{where} \quad P_A(\vec{r}) = \prod_{B \neq A} s_A(\mu_{AB}). \tag{4}$$

The step function (3) can be replaced by a continuous, monotonically decreasing function in the range $(-1, 1)$, fulfilling the requirements $s(-1) = 1$ and $s(+1) = 0$ in order to define the fuzzy Voronoi cells. Becke suggested the simple polynomial function

$$s_A^k(\mu_{AB}) = \frac{1}{2}[1 - f_k(\mu_{AB})], \tag{5}$$

where $f_1(\mu) = \frac{3}{2}\mu - \frac{1}{2}\mu^3$ and $f_k(\mu) = f[f_{k-1}(\mu)]$. The integer k is known as the *stiffness* parameter and controls the shape of the cutoff profile. The larger the value of k the steeper is the cutoff profile.

The fuzzy Voronoi cells thus defined do not account for the different atomic sizes in heteronuclear systems, as the faces of the Voronoi cells exactly bisect the internuclear axis between the two neighboring atoms. A shifted cutoff profile can be obtained with the cutoff function (5) using as an argument the following transformed coordinate:

$$\nu_{AB} = \mu_{AB} + a_{AB}\left(1 - \mu_{AB}^2\right), \tag{6}$$

where $-1/2 \le a_{AB} \le 1/2$ in order to ensure that $-1 \le \nu_{AB} \le 1$. The position of the cell boundary is now controlled by the value of the parameter a_{AB}, which in the new coordinates is given by the condition $\nu_{AB} = 0$. At the point where the shifted cell boundary intersects the interatomic plane, one has

$$\mu_{AB} = \frac{r_A - r_B}{r_A + r_B}. \tag{7}$$

Becke suggested to relate the ratio r_A/r_B with the relative size of the atoms A and B making use of some set of reference *fixed* atomic radii R_A^0 and R_B^0 as

$$\frac{r_A}{r_B} = \frac{R_A^0}{R_B^0} = \chi_{AB}. \tag{8}$$

In his original paper, Becke used the set of empirical atomic radii of Bragg and Slater[14] and the value $k = 3$, on the basis of a better performance of his numerical integration scheme.

Even though Becke's atoms are only mathematical constructs, we have shown that, with some special provisions, they can be readily used as atoms in molecules models, from which chemically sound values of atomic populations, bond orders, or one and two-center energy components can be obtained. However, the use of a set of *fixed* atomic radii is a clear limitation of this model, basically because same atoms are treated on equal footing in different chemical environments, i.e., partial ionic character of atoms may not be properly taken into account.

One way to sort out this problem is to use an internal criterion to determine the position of the cell boundaries between all neighboring atoms. Thus, instead of a set of fixed radii, one needs to establish a strategy to define the set of χ_{AB} values, where A and B represent any pair of atoms sharing a cell boundary. In Ref. 11, some of us proposed as a criterion to use the position of the extremum (typically a minimum) of the density along the internuclear axis connecting every two neighboring atoms to locate their cell boundary. That scheme was later referred to as Becke-ρ in Ref. 15. A similar scheme making explicit use of the bond critical points of the density was also proposed by Francisco *et al.*[16] Not surprisingly, the Becke-ρ scheme typically yielded electron populations, bond orders, and energy components close to those obtained by applying the disjoint QTAIM atoms.[15–17]

There is, however, a limitation of the original formulation of Becke-ρ, that also applies to the original Becke scheme itself. We just need to recall that the transformation (6) does not permit an arbitrary shift of the cell boundary because if $|a_{AB}| > \frac{1}{2}$ the transformation, and hence the resulting cutoff profile, would not be monotonic. The maximum possible shift of the cell boundary occurs when the atomic radii of the two atoms differ by a factor of $1 + \sqrt{2}$ (ca. 2.4). Such ratio may be easily reached when involving pairs of atoms of most different size and electronegativity, which renders a limitation of the method when used as an alternative to QTAIM. It is fair to note that such limitation in the shift of the cell boundaries was already noted by Becke in the Appendix of Ref. 13, which apparently has passed unnoticed so far.

In order to be able to fully accommodate the criterion behind the original definition of the Becke-ρ scheme, we propose here the following alternative transformation:

$$\nu'_{AB} = \frac{1 + \mu_{AB} - \chi_{AB}(1 - \mu_{AB})}{1 + \mu_{AB} + \chi_{AB}(1 - \mu_{AB})}. \tag{9}$$

One can easily verify that $\nu'_{AB} = \mu_{AB}$ at the limit values of $\mu_{AB} = -1$ and $\mu_{AB} = 1$, and also if $\chi_{AB} = 1$, i.e., in the homonuclear case.

The alternative new transformation is monotonic for any value of χ_{AB}, so it can be applied to shift the cell boundary to any position along the internuclear axis. It is worth to note that the transformations (6) and (9) lead to different cutoff profiles, even within the limits of applicability of the former. The differences are minimal yet the new definition exhibits a somewhat less sharp profile, as shown in Figure S1 of the supplementary material.[18]

We have implemented this new scheme, henceforth, topological fuzzy Voronoi cells (TFVC), in our APOST-3D program.[19] For TFVC, we have adopted a stiffness parameter of $k = 4$. We have observed a better overall agreement with QTAIM atomic populations than for $k = 3$ without a significant loss of numerical accuracy in the overall integration. The agreement with the QTAIM bond orders or delocalization indices (DIs) also improves with $k = 4$, in particular for non-bonded atoms. The observation that in benzene the DI between carbon atoms in para position is larger than in meta position[20] is better reproduced with this TFVC model.

In our implementation, a rather modest grid of 40 radial and 146 angular points per atom typically yields the total

number of electrons with an error of ca. 10^{-3}. For the numerical integration of other quantities such as one-electron energy components, we use a larger grid of 70×434 points per atom. In a most recent implementation of QTAIM integrations,[21] the author uses up to 0.87M grid points per atom for a typical run. The reduction of the atomic grid size in fuzzy approaches is dramatic due to the lack of strict atomic boundaries. These methods can integrate all centers at once and scale linearly with the total number grid points, essentially for any level of theory. Most recently, delocalization indices for a set of fullerenes were determined using Becke-ρ.[22] We have included in the supplementary material[18] a comparative analysis of the accuracy of TFVC and QTAIM for a challenging endohedral complex.

Another advantage of the TFVC scheme is that it is flexible enough to either ignore the presence of spurious NNAs, as described above, or to readily incorporate them if necessary by defining extra fuzzy Voronoi cells centered on the position of the attractors. In the supplementary material,[18] we provide illustrative examples of the TFVC method with NNAs.

Finally, it is worth to note that, strictly speaking, there are no boundaries in the fuzzy Voronoi cells. Thus far we have referred to cell boundary between atoms A and B to the intersecting plane for which $\mu_{AB} = 0$. Hence, each pair of Voronoi cells (atoms), whether they are chemically bonded or not, has a formal boundary, whose location must be established. In the Becke-ρ method,[11] if two atoms were too far to be considered as *neighbors* (the distance was larger than twice the sum of their fixed atomic radii), the ratio given in Eq. (8) was used to locate the fuzzy cell boundary. This leads to some inconsistencies when dealing with intermolecular interactions of bond dissociation curves, where atoms sharing cell boundaries may be too far away to be neighbors. In TFVC, we have adopted another criterion that also avoids any use of arbitrary fixed atomic radii. Thus, a given pair of atoms are not considered as neighbors if their midpoint is closer to a third nucleus or attractor of the molecular system. In that case, the formal cell boundary is placed at the midpoint, that is, a $\chi_{AB} = 1$ value is taken.

We have computed with TFVC the atomic populations and delocalization indices of the molecular set described in Ref. 2, which features a number of hydrides as well as prototypical organic molecules. The hydrides have been chosen because the limits of applicability of the original Becke-ρ scheme are easily reached when one of the involved atoms is a hydrogen. The wavefunctions and densities have been obtained at the fully optimized Hartree-Fock and B3LYP levels of theory with the cc-pVTZ basis set. The agreement between the TFVC and QTAIM atomic charges at B3LYP level for the molecular set is very good, as indicated by the values of R^2 and the slopes and intercepts of Fig. 1. Homonuclear diatomics and equivalent atoms have been excluded in the comparison. The results at the Hartree-Fock level are very similar and are not discussed. The mean unsigned error between TFVC and QTAIM charges is below 0.12, and the larger discrepancies are observed for hydrides involving 3rd period atoms with electronegativity similar to that of the H atom. The agreement between the TFVC and QTAIM delocalization indices is also strikingly good (see Fig. 2). The mean

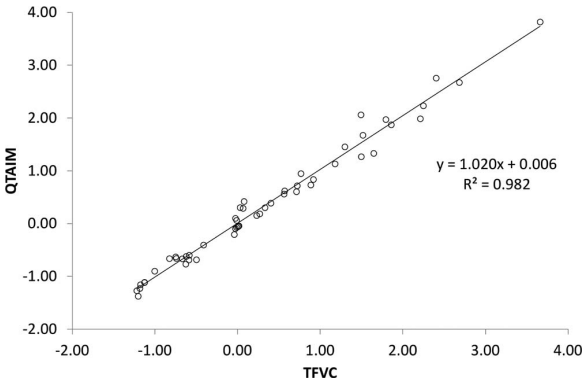

FIG. 1. QTAIM vs TFVC atomic charges for the molecular set.

unsigned error is 0.06 and the largest one is just 0.25. The corresponding plots using $k = 3$ are included in the supplementary material.[18]

An obvious physical limitation of the TFVC method is that the cell boundaries are described by planes, whereas in QTAIM they are typically curved. However, neither the overlapping character of the boundaries nor their curvature seems to be critical for TFVC integrations to be close to QTAIM. In our experience, it is when the electron density on the interatomic region is very flat (and large) that the QTAIM and TFVC values are more likely to diverge (see the supplementary material[18] for a deeper analysis).

Finally, for illustration purposes we have also compared the behavior of the TFVC scheme in the decomposition of the Hartree-Fock molecular energy.[3,16] Table I gathers the different contributions (namely, kinetic, electron-nuclear repulsion, Coulomb, and exchange) to the one and two-center energy components for hydrogen fluoride at the RHF/cc-pVTZ level of theory. The QTAIM results have been obtained with the program AIMALL.[23] The overall accuracy of the numerical integration, expressed as the difference between the total energy and the sum of the one- and two-center terms separately, is similar for all methods. The two-electron numerical integrations in TFVC and Becke-ρ schemes have been carried out using solely a pair of rotated grids of 40×146 points per

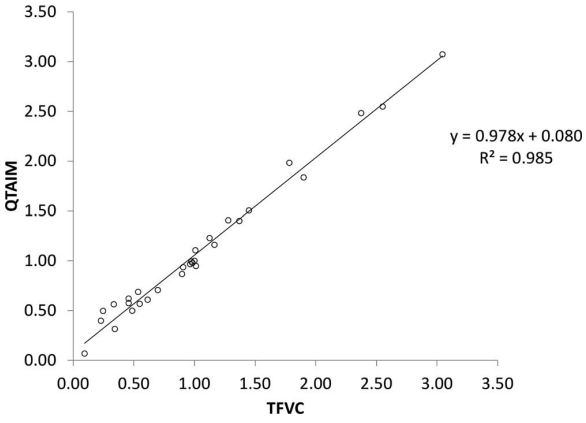

FIG. 2. QTAIM vs TFVC delocalization indices for the molecular set.

TABLE I. Atomic populations, DI, integrated laplacian, and atomic and diatomic energy terms (in a.u.) for HF molecule computed at the HF/cc-pVTZ level of theory. T, V_{en}, V_{Coul}, V_x, E_{tot}, and δE_{tot} are the kinetic, electron-nuclear attraction, Coulomb repulsion, exchange, total energies, and integration error, respectively.

Term	QTAIM	TFVC	Becke-ρ
N(H)	0.219	0.237	0.582
N(F)	9.778	9.762	9.418
DI (H,F)	0.394	0.425	0.913
Lapl.(H)	0.0001	− 0.0102	− 0.2002
Lapl.(F)	0.0000	0.0168	0.2009
T(H)	0.2382	0.2464	0.4223
T(F)	99.8173	99.8090	99.6331
V_{en}(H)	− 0.4417	− 0.4641	− 0.8911
V_{en}(F)	− 243.7306	− 243.6379	− 241.8384
V_{en}(H,F)	− 6.7086	− 6.7794	− 8.1515
V_{Coul}(H)	0.0322	0.0361	0.1705
V_{Coul}(F)	54.8719	54.7666	52.8938
V_{Coul}(H,F)	1.0091	1.1062	2.8473
V_x(H)	− 0.0152	− 0.0168	− 0.0709
V_x(F)	− 10.3095	− 10.2964	− 10.0606
V_x(H,F)	− 0.1253	− 0.1369	− 0.3179
E_{tot}(H)	− 0.1865	− 0.1985	− 0.3691
E_{tot}(F)	− 99.3509	− 99.3587	− 99.3721
E_{tot}(H,F)	− 0.5206	− 0.5059	− 0.3179
δE_{tot}	0.88×10^{-3}	0.66×10^{-3}	0.20×10^{-3}

atom, as described in Ref. 12. On the other hand, the value of the Laplacian of the density integrated over the atomic domain is a typical measure of the fulfillment of the zero-flux condition within QTAIM. The value obtained with TFVC (ca. 10^{-2}) is one order of magnitude smaller than that obtained with Becke-ρ scheme. This value is somewhat too large compared with QTAIM, however, it does not seem to influence too much the values of the different one and two-center contributions to the energy. The TFVC energy contributions are much closer to QTAIM that to the Becke-ρ ones, especially the purely electrostatic ones. It is worth to note that for the sake of better comparison, the Becke-ρ values on Table I have been obtained using $k = 4$. Thus, the main difference between TFVC and Becke-ρ in this case is the proper location of the fuzzy cell boundary.

In conclusion, the TFVC method represents fast and simple atoms in molecules scheme that can be routinely used to extract chemical information from large-scale *ab initio* calculations. The method presented here can be regarded as a general purpose computationally more efficient alternative to Bader's QTAIM.

Financial help has been furnished by the Spanish MICINN Project No. CTQ2011-23441/BQU, the FEDER UNGI08-4E-003, and the Generalitat de Catalunya SGR528 grants. Support from the Xarxa de Referència en Química Teòrica i Computacional is also acknowledged. E.R.-C. acknowledges support from the Spanish FPU program (Grant No. AP2008-01231).

[1] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, Oxford, 1990).
[2] X. Fradera, M. A. Austen, and R. F. Bader, J. Phys. Chem. A **103**, 304 (1999).
[3] P. Salvador, M. Duran, and I. Mayer, J. Chem. Phys. **115**, 1153 (2001).
[4] M. A. Blanco, A. M. Pendás, and E. Francisco, J. Chem. Theory Comput. **1**, 1096 (2005).
[5] C. Gatti, Z. Kristallogr. **220**, 399 (2005).
[6] J. Friis, G. K. H. Madsen, F. K. Larsen, B. Jiang, K. Marthinsen, and R. Holmestad, J. Chem. Phys. **119**, 11359 (2003).
[7] A. M. Pendás, M. A. Blanco, A. Costales, P. Mori-Sánchez, and V. Luana, Phys. Rev. Lett. **83**, 1930 (1999).
[8] Q. K. Timerghazin, I. Rizvi, and G. H. Peslherbe, J. Phys. Chem. A **115**, 13201 (2011).
[9] D. Chesnut, Heteroat. Chem. **13**, 53 (2002).
[10] D. R. Alcoba, L. Lain, A. Torre, and R. C. Bochicchio, Chem. Phys. Lett. **407**, 379 (2005).
[11] I. Mayer and P. Salvador, Chem. Phys. Lett. **383**, 368 (2004).
[12] P. Salvador and I. Mayer, J. Phys. Chem. **120**, 5046 (2004).
[13] A. Becke, J. Chem. Phys. **88**, 2547 (1988).
[14] J. C. Slater, J. Chem. Phys. **41**, 3199 (1964).
[15] E. Matito, M. Solà, P. Salvador, and M. Duran, Faraday Discuss. **135**, 325 (2006).
[16] E. Francisco, A. M. Pendás, and M. A. Blanco, J. Chem. Theory Comput. **2**, 90 (2006).
[17] W. Heyndrickx, P. Salvador, P. Bultinck, M. Solà, and E. Matito, J. Comput. Chem. **32**, 386 (2011).
[18] See supplementary material at http://dx.doi.org/10.1063/1.4818751 for additional data such as integration accuracy of TFVC and use of the method with NNA's.
[19] P. Salvador and E. Ramos-Cordoba, "APOST-3D," University of Girona, Spain, 2013.
[20] J. Poater, X. Fradera, M. Duran, and M. Solà, Chem.-Eur. J. **9**, 400 (2003).
[21] P. M. Polestshuk, J. Comput. Chem. **34**, 206 (2013).
[22] M. Garcia-Borràs, S. Osuna, M. Swart, J. M. Luis, and M. Solà, "Maximum aromaticity as a guiding principle for the most suitable hosting cages in endohedral metallofullerenes," Angew. Chem. (published online).
[23] AIMAll (Version 13.05.06), Todd A. Keith, TK Gristmill Software, Overland Park KS, USA, 2013, see aim.tkgristmill.com.

SUPPORTING INFORMATION

An approximation to Bader's topological atom

Pedro Salvador, Eloy Ramos-Cordoba[1, a)]

Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, 17071 Girona, Spain

Contents

[a)] pedro.salvador@udg.edu

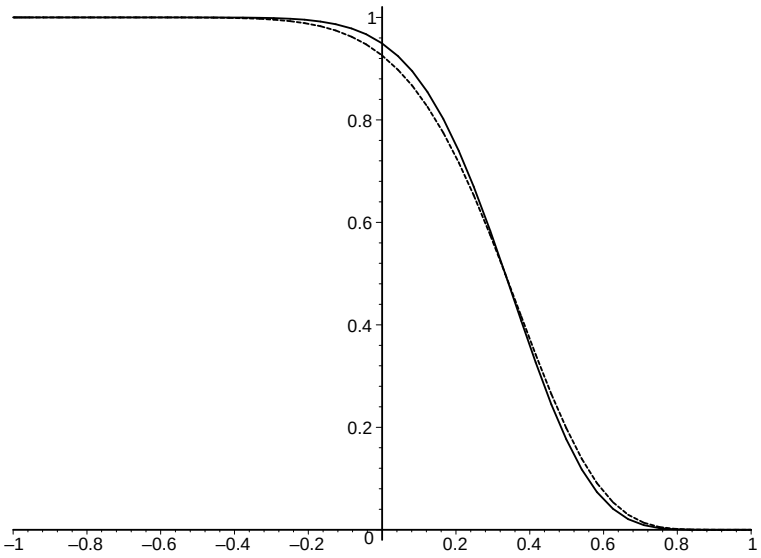

FIG. S1. Cutoff profiles with $k = 3$ and $\chi_{AB} = 2$ for the original Becke (solid) and the alternative TFVC (dash) formulation.

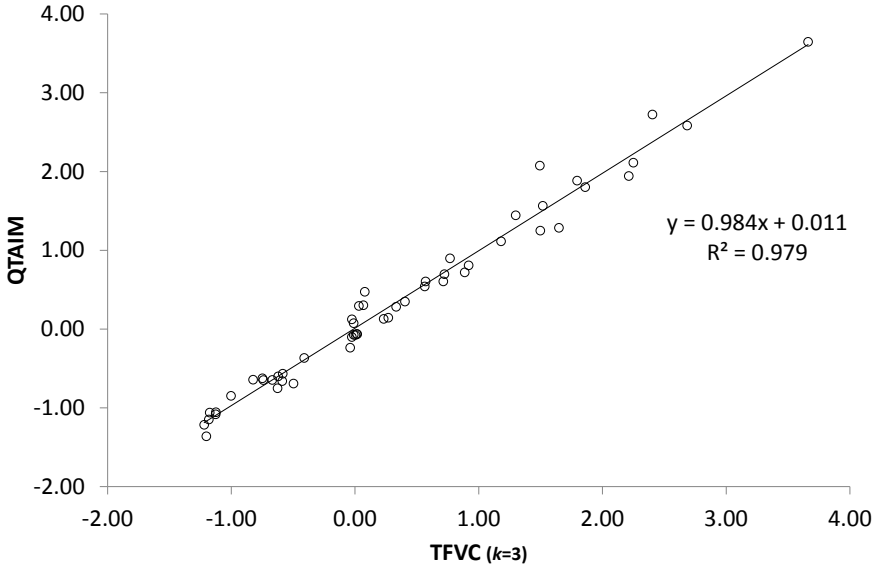

FIG. S2. QTAIM vs TFVC with stiffness parameter $k = 3$ atomic charges for the molecular set.

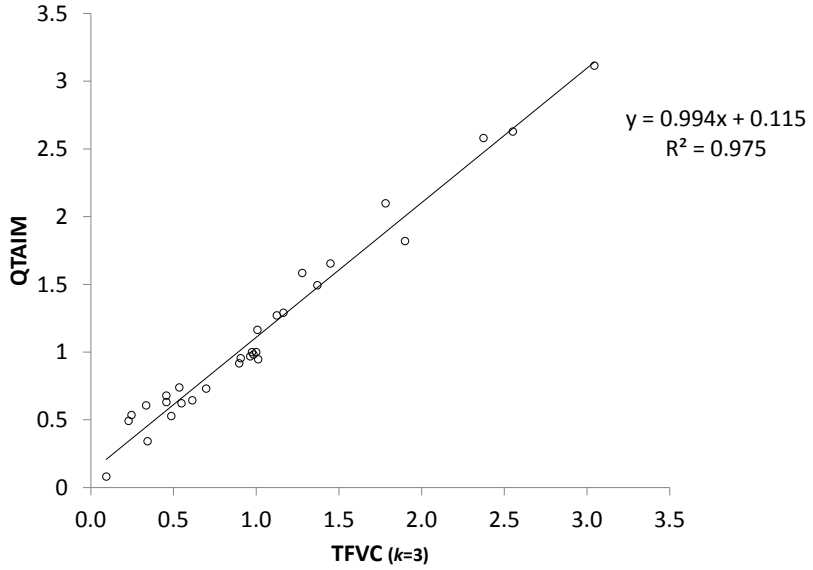

FIG. S3. QTAIM vs TFVC with stiffness parameter $k = 3$ delocalization indices for the molecular set.

A. Performance of TFVC for non-nuclear attractors

The TFVC method can be easily adapted to deal with non-nuclear attractors (NNAs) of the density. We present here some results for acetylene (for which the particular combination of level of theory and basis set lead to the formation of a spurious NNA in the midpoint of the C-C bond) and the dimeric Mg(I) compound of Fig.S4, that exhibits a genuine NNA between the two Mg atoms. The atomic coordinates have been taken from Ref. S1, and the large diisopropylphenyl groups have been replaced by methyl groups. The wavefunction has been obtained at the BP86/6-31+G** level of theory.

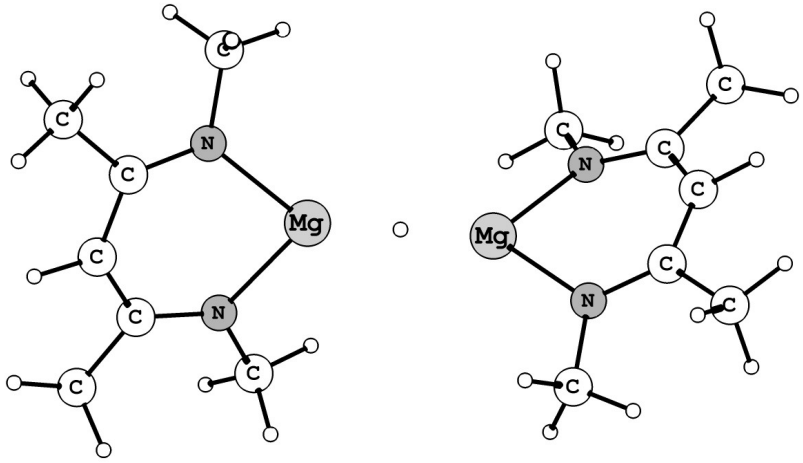

FIG. S4. Model dimeric Mg(I) compound with the NNA between the Mg centers

On Table S1 one can see that the electron population of the NNA on acetylene is clearly underestimated with the TFVC method with respect to QTAIM by more than 1 electron. The populations of the H atoms are quite similar. The density on the interatomic region between the two C atoms is extremely flat, so the value of the density at the NNA is similar to that of the bcp connecting the C atom. As a result, small displacements of the position of the fuzzy Voronoi boundary are translated into important changes on populations of the two basins. Indeed, with a shift of 0.07Å on the location of the fuzzy boundary the TFVC population of the NNA increases to 1.60. It is worth to note that the zero-flux surface between the NNA and the C atom does not deviates too much from planarity, which indicates that the shape of the TFVC and QTAIM basins is not too different. In the case of the dimeric Mg(I) compund, the differences between QTAIM and TFVC populations of the genuine NNA and neighboring atoms are still significant.

Nevertheless, there is no fundamental difference in describing an atom or a NNA with TFVC. The observed deviations should not be associated to the nature of the attractor but to the shape of the density in the interatomic regions. In fact, after careful analysis we have come to the conclusion that neither the overlapping character of the boundaries, nor their curvature are critical for TFVC populations to be close to QTAIM. In Figures S5 and S6 we compare the shape of the atomic basins with QTAIM and TFVC for H_2O and H_2S systems described at the RHF/cc-pVTZ levels of theory. The H_2S correspond to the worst case of our molecular set.

In the case of H_2O the curvature of the zero-flux surface is very pronounced so the shape of the QTAIM and TFVC atoms are quite different. Yet, the integrated electron populations (see TableS1) are strikingly close. The same occurs with hydrogen fluoride, for instance. On the other hand, in H_2S the shape of the QTAIM and TFVC boundaries are quite alike but the differences in the integrated quantities are significant (almost 1 e for the S atom), similar to those observed for the NNA of acetylene.

In conclusion, our results indicate that when the electron density on interatomic region is very flat (and large), small deviations on the location of the fuzzy Voronoi boundaries can produce significant changes on the integrated quantites, and thus the QTAIM and TFVC are more likely to diverge.

TABLE S1. Atomic populations of selected atoms and NNAs with QTAIM and TFVC methods.

Molecule/Atom	QTAIM	TFVC
Acetylene/C	5.274	5.848
Acetylene/H	0.861	0.930
Acetylene/NNA	1.727	0.442
Mg dimer/Mg	10.695	11.070
Mg dimer/NNA	0.825	0.154
H_2O/O	9.272	9.258
H_2S/S	16.635	17.502

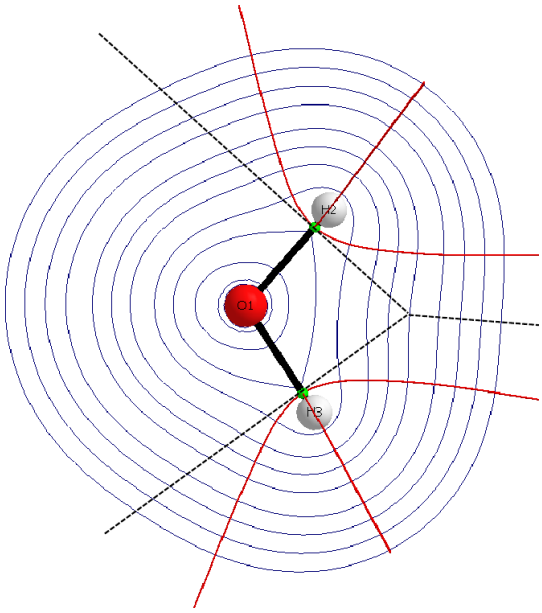

FIG. S5. Voronoi cell boundary (dashed line) and interatomic surface paths (red lines) for H_2O in the molecular plane.

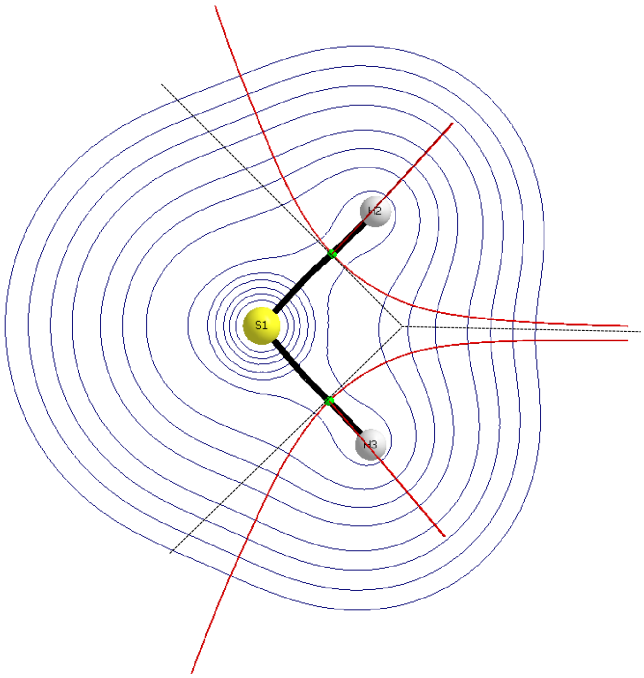

FIG. S6. Voronoi cell boundary (dashed line) and interatomic surface paths (red lines) for H_2S in the molecular plane.

B. Convergence of TFVC numerical integrations

The goal of the accompanying manuscript is to present a new, more flexible, atomic definition for an atom in a molecule that mimics some of the features of Bader's QTAIM. It is not about an efficient implementation of an existing scheme.The method is claimed to be efficient because of its simplicity, not because of the way it has been implemented computationally, which at present is still far from optimum. Thus, in our implementation of TFVC we use the same grid size and shape for all atoms int he system, and the numerical integrations of all atoms are performed at once, contrary to most QTAIM implementations when one typically carries out the basin integrations atom by atom.

In Table S2 we gather some data about the change on the atomic populations with the size of the atomic grid used for the numerical integrations. We have chosen for illustration a endohedral complex of the helium atom and adamantane, which allegedly is a challenging system for topological analysis. We have collected the partial atomic charges of the He and C atoms (C^t atoms exhibit a single C-H bond, the C^s present two), overall accuracy of the numerical integration of the electron density and CPU times. We include also data from Ref. 22. Note that one can hardly compare CPU times with other programs that perform QTAIM analysis because they probably involve integrations of a number of different functions and in this case they ran on a different architecture.

We feel more relevant to discuss the number of grid points per atom used for the numerical integration. In Ref. 22 the author used almost a million integrations points, for an overall accuracy of 2.0E-03. With an even larger grid and special provisions an excellent accuracy of 7.5E-05 was finally achieved. In our case we use for a typical run a combination of 40x146=5840 integrations points per atom to perform the integrations of the density, MOs and two-electron integrations. Only in the case of one-electron energy contributions we use up to 70x434=30380 integration points to achieve better accuracy on the energy contribution (see Table I of the accompanying manuscript). The reduction is dramatic but merely because the shape of the atomic basins is considerably less cumbersome in TFVC compared with QTAIM, so less number of integration points are needed to achieve sufficient accuracy.

The data on Table S2 show an almost monotonic improvement of the overall integration accuracy with the atomic grid size, which is not alway the case. The He@adamantane system shows indeed some difficulties of the TFVC to achieve an acceptable accuracy with

the typical 40x146 atomic grid. For this challenging system a 50x230 grid is necessary. The overall accuracy could probably be further improved by using different grid sizes for different atoms. On the other hand, the CPU times shows that the method scales rather linearly with the atomic grid size.

TABLE S2. Partial atomic charges for He, C^t and C^s atoms of He@adamantane. Sum indicates the overall integration error on the sum of electrons populations. CPU time in minutes on a single-processor Intel Xeon CPU X5650 @ 2.67GHz

Atomic grid	He	C^t	C^s	Sum	CPU (min)
30x110	-0.2653	0.2417	0.1562	-5.4E-02	6.5
40x110	-0.2694	0.2396	0.1564	-5.9E-02	9.5
40x146	-0.2779	0.2447	0.1601	-2.1E-02	12.5
50x146	-0.2760	0.2457	0.1592	-2.4E-02	15
50x230	-0.2709	0.2494	0.1601	3.7E-03	28
50x434	-0.2681	0.2483	0.1599	-1.8E-03	76
70x434	-0.2680	0.2483	0.1597	-2.5E-3	152
TWOE-LLG[a]	-0.0988	0.1046	0.0628	2.0E-03	113
TWOE-GLG[a]	-0.0989	0.1040	0.0630	7.5E-05	653

[a] Data from ref. 22. LLG used 0.87M grid points for all atoms. GLG used 2.3M and 4.2M grid points for He and C atoms, respectively. CPU times on an Intel Core i7, 1.8 GHz laptop dual-core processor.

REFERENCES

S1. J. Overgaard, C. Jones, A. Stasch, and B. B. Iversen, J. Amer. Chem. Soc. **131**, 4208–4209 (2009)

Chapter 4

On the proper definition of local spin

4.1 Toward a unique definition of the local spin

Ramos-Cordoba, E.; Matito, E.; Mayer, I.; Salvador, P. “Toward a Unique Definition of the Local Spin”. *Journal of Chemical Theory and Computation*. 2012, 8, 1270–1279

http://dx.doi.org/10.1021/ct300050c

http://pubs.acs.org/doi/abs/10.1021/ct300050c

Received: January 26, 2012

Published: February 24, 2012

Abstract

In this work, we demonstrate that there is a continuum of different formulations for the decomposition of [$\hat{S}2$] that fulfills all physical requirements imposed to date. We introduce a new criterion based upon the behavior of single-electron systems to fix the value of the parameter defining that continuum, and thus we put forward a new general formula applicable for both single-determinant and correlated wave functions. The numerical implementation has been carried out in the three-dimensional physical space for several atomic definitions. A series of representative closed-shell and open-shell systems have been used to illustrate the performance of this new decomposition scheme against other existing approaches. Unlike other decompositions of [$\hat{S}2$], the new scheme provides very small local-spin values for genuine diamagnetic molecules treated with correlated wave functions, in conformity with the physical expectations.

4.2 Local spins: improved Hilbert-space analysis

PCCP

Cite this: *Phys. Chem. Chem. Phys.*, 2012, **14**, 15291–15298

www.rsc.org/pccp

PAPER

Local spins: improved Hilbert-space analysis

Eloy Ramos-Cordoba,[a] Eduard Matito,[a] Pedro Salvador*[a] and István Mayer[b]

Received 29th May 2012, Accepted 17th September 2012
DOI: 10.1039/c2cp42513k

The decomposition of $\langle \hat{S}^2 \rangle$ for a general wave function has been carried out in the framework of the Hilbert-space analysis. The one and two-center components fulfill all physical requirements imposed to date. An inherent ambiguity of the Hilbert-space decomposition of a two-electron quantity, in particular using a Mulliken-type scheme, is also discussed in detail. The formalism of effective atomic densities has allowed us to derive in a simple manner appropriate expressions for the decomposition of $\langle \hat{S}^2 \rangle$ in the framework of Hilbert space analysis that are consistent with Mulliken population analysis and related quantities. Using a particular mapping we have derived the Hilbert-space expressions also in the framework of Löwdin population analysis in a straightforward manner. The numerical results obtained with the latter formalism have proved to be more robust and reliable.

1 Introduction

The concept of local spin emerges in a quite natural fashion when describing the electronic structure of systems with diradical character such as non-Kekulé molecules or transition state structures of chemical reactions. Heisenberg Hamiltonian models also invoke the concept of local spin in order to assess the nature of spin–spin interactions between magnetic centers. Often, the spin properties of a molecule can be characterized by the spin density. There are, however, cases where the overall system is a singlet (where there is no spin density), but for which the existence of some local spin is assumed. In the last few years there has been a growing interest in recovering local spins from the analysis of the wave function of *ab initio* calculations.[1–15] Different schemes have been proposed in the literature, most of which are rooted in the decomposition of the expectation value of the spin squared operator into atomic and diatomic contributions, for both single-determinant and correlated wave functions. Because the partitioning of the single physical quantity $\langle \hat{S}^2 \rangle$, which in the case of singlet wave functions is *zero*, into components is not unique, a number of physical requirements[4,9,15] have been introduced.

(i) One should get no spins whatever for covalent systems described by a closed-shell RHF wave function using doubly-filled orbitals.

(ii) If the wave function is properly dissociating, then the asymptotic values of the atomic spins obtained for the atoms at large distances should coincide with the corresponding values of the free atoms.

(iii) In an open-shell system the overall $\langle \hat{S}^2 \rangle$ does not depend on the actual $\hat{S}_z$ projection of the electronic state (multiplet) considered, so one may request to have $\langle \hat{S}^2 \rangle$ components that do not depend on $\hat{S}_z$ either.

(iv) No two-center terms should appear in the case of single-electron systems (or ROHF systems with a single unpaired electron).

In a previous paper[15] we showed that the following general expression

$$\langle \hat{S}^2 \rangle = \tfrac{3}{4}\int u(\vec{r}_1)\,\mathrm{d}\vec{r}_1 + \tfrac{1}{2}\int\int[\Gamma(\vec{r}_1,\vec{r}_2) - \tfrac{1}{2}\rho^{\mathrm{s}}(\vec{r}_1;\vec{r}_2)\rho^{\mathrm{s}}(\vec{r}_2;\vec{r}_1)]\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 - \tfrac{1}{2}\int\int[\Gamma(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1) - \tfrac{1}{2}\rho^{\mathrm{s}}(\vec{r}_1;\vec{r}_1)\rho^{\mathrm{s}}(\vec{r}_2;\vec{r}_2)]\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 \quad (1)$$

is the natural starting point to derive atomic and diatomic components of $\langle \hat{S}^2 \rangle$ that satisfy requirements (i) to (iv). This equation is written in terms of the density of effectively unpaired electrons, $u(\vec{r})$, defined by Takatsuka *et al.*[16] as

$$u(\vec{r}) = 2\rho(\vec{r}) - \int\rho(\vec{r};\vec{r}')\rho(\vec{r}';\vec{r})\mathrm{d}\vec{r}', \quad (2)$$

the spin-density matrix

$$\rho^{\mathrm{s}}(\vec{r};\vec{r}') = \rho^{\alpha}(\vec{r};\vec{r}') - \rho^{\beta}(\vec{r};\vec{r}'), \quad (3)$$

and the spin-less cumulant of the second order density matrix, $\Gamma(\vec{r}_1,\vec{r}_2;\ \vec{r}'_1,\vec{r}'_2)$, which vanishes for single-determinant wave functions and can be defined as the sum of the usual (spin-dependent) cumulants as

$$\Gamma(\vec{r}_1,\vec{r}_2;\ \vec{r}'_1,\vec{r}'_2) = \sum_{\sigma,\sigma'}\Gamma^{\sigma\sigma'\sigma\sigma'}(\vec{r}_1,\vec{r}_2;\ \vec{r}'_1,\vec{r}'_2) \quad (4)$$

[a] *Department of Chemistry and Institute of Computational Chemistry, University of Girona, 17071 Girona, Spain. E-mail: Pedro.Salvador@udg.edu*
[b] *Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary*

where

$$\Gamma^{\sigma\sigma'\sigma\sigma'}(\vec{r}_1,\vec{r}_2;\ \vec{r}'_1,\vec{r}'_2) = \rho_2^{\sigma\sigma'\sigma\sigma'}(\vec{r}_1,\vec{r}_2;\ \vec{r}'_1,\vec{r}'_2) - \rho^{\sigma}(\vec{r}_1,\vec{r}'_1)\rho^{\sigma'}(\vec{r}_2,\vec{r}'_2) - \delta_{\sigma\sigma'}[\rho^{\sigma}(\vec{r}_1,\vec{r}'_2)\rho^{\sigma}(\vec{r}_2,\vec{r}'_1)]. \quad (5)$$

In ref. 15 we obtained one and two-center contributions for a general wave function in the framework of the 3D-space analysis, *i.e.*, for "fuzzy atoms"[17] and Bader's atomic domains.[18]

In this paper we wish to undertake the decomposition in the context of the so-called "Hilbert-space analysis".[19] The motivation is twofold: first, aside its conceptual relevance, the Hilbert-space decomposition does not require atomic numerical integrations, in contrast to the 3D-space analysis; thus it is exact (is free of the numerical errors of that integration). Also, the significant reduction in the computational cost of the decomposition may be relevant for very large systems, especially as compared to the 3D-space methods with complicated atomic basins of Bader's analysis. However, there is an apparent ambiguity in decomposing two-electron quantities in the framework of Hilbert-space analysis (in particular for the Mulliken-type scheme), which to date has not received due attention. In this paper we also wish to analyze in more detail this problem, which is particularly relevant in the case of the decomposition of $\langle \hat{S}^2 \rangle$. That ambiguity will be exposed in the next section. Then, we will briefly describe the formalism of effective atomic densities,[20] which will allow us to derive in a straightforward manner the most appropriate expressions for the decomposition of $\langle \hat{S}^2 \rangle$ in the framework of Hilbert space analysis. Finally, some numerical results at the correlated level will be presented and discussed.

2 Alternative summation schemes in the Hilbert-space analysis

The decomposition of physical quantities into atomic and diatomic contributions is rooted on the identification of an atom within the molecule. Since practical quantum chemistry mostly uses atom-centered basis sets, the atom may be identified with its nucleus and the subspace spawned by the set of atomic basis functions centered on it. The simplest example of application of such Hilbert-space analysis is Mulliken population analysis,[21] perhaps the most familiar method to determine the number of electrons associated with an atom. Mulliken's gross population of atom A is defined as:

$$N_A = \sum_{\mu\in A}\sum_{\nu} D_{\mu\nu}S_{\nu\mu} = \sum_{\mu\in A}(\boldsymbol{DS})_{\mu\mu}, \quad (6)$$

where the notation $\mu \in A$ indicates that the summation runs over all atomic basis functions centered on atom A. We recall in this context that matrix $\boldsymbol{DS}$ is the proper finite basis representation of the first-order density matrix if an overlapping basis set ($\boldsymbol{S} \neq \boldsymbol{I}$) is used.[22]

In a similar manner, the Mayer–Wiberg (closed-shell for simplicity) bond order,[23] B_{AB}, between atoms A and B is defined as

$$B_{AB} = \sum_{\mu\in A}\sum_{\sigma\in B}(\boldsymbol{DS})_{\mu\sigma}(\boldsymbol{DS})_{\sigma\mu}. \quad (7)$$

Inspecting the expression in eqn (6), one can see that the overlap integrals enter it in a somewhat non-symmetric manner: one of the subscripts (μ) is serving for subdividing the quantity into atomic contributions, while another (ν) is a "dummy" index, for which summation over the whole basis is performed—it is used to form the matrix-product $\boldsymbol{DS}$. This difference may be connected with the fact that for overlapping basis sets matrix $\boldsymbol{DS}$ is twice the projection matrix performing the projection of any vector $\boldsymbol{d}$ of LCAO coefficients on the subspace of the occupied molecular orbitals as $\boldsymbol{DSd}$.[22] The same distinction appears also in eqn (7) of the Mayer–Wiberg bond order. In the case of real orbitals, one could get exactly the same Mulliken atomic populations also in the form $\sum_{\mu\in A}(\boldsymbol{SD})_{\mu\mu}$, *i.e.*, by using matrix $\boldsymbol{SD}$ which performs the analogous projection of the *row-vectors* $\boldsymbol{d}^\dagger$ as $\boldsymbol{d}^\dagger\boldsymbol{SD}$. While in the first case the systematization of the terms according to the individual atoms corresponds to the subscript coming from the "ket" part of the overlap integral, in the second one it corresponds to the subscript coming from its "bra" part. It seems logical to stick to one of these possibilities (we prefer the first one), and use it in all types of analyses. Thus the splitting of the terms in the expression of the bond order index eqn (7) corresponds to the subscripts of the overlap integrals coming from the "kets".

In principle, if an expression contains products with two overlap matrices, then a subdivision into atomic and diatomic contributions by taking one subscript from "bra" and another from "ket" is also possible. In the case of the bond order, that leads to a modified definition of the bond order index[24] as

$$B'_{AB} = \sum_{\mu\in A}\sum_{\nu\in B}(\boldsymbol{SDS})_{\mu\nu}\boldsymbol{D}_{\nu\mu} \quad (8)$$

As the bond order is a component of the integral of the exchange density, formally both definitions could be acceptable: they represent different decompositions of that integral into a sum of one- and two-center contributions. However, there is a serious argument favoring the definition of eqn (7). The modified, non-symmetric, definition eqn (8) gives results that are much less "chemical" than those given by the original one: it cannot, for instance, recover the integer values for first-row diatomics (*e.g.*, 3 for N_2) if a minimal basis set is used, as does the original definition of eqn (7). Another argument against such type of "bra"–"ket" mixing is the high degree or arbitrariness that would be introduced in the case of *e.g.* Generalized Population Analysis,[25] typically used to detect patterns of multicenter bonding, where the expressions may contain three, four or more overlap matrices. Furthermore, it has been shown[26] that one can introduce a particular mapping (see Appendix) between the atomic overlap matrices of the atomic orbitals and the conventional overlap matrix that permits finding a one-to-one correspondence between the Hilbert-space and the more general 3D-space analyses expressions of quantities like bond orders, atomic valences or energy components. Such a general mapping is not possible for expressions involving subdivision of the terms according to both "bra" and "ket" subscripts.

These considerations are of interest in the present context because in the recent paper by Alcoba *et al.*[9] a decomposition

of $\langle \hat{S}^2 \rangle$ is performed in a manner that one index of the overlap matrix is assigned according to the term coming from the "bra" and another coming from the "ket". (These authors distinguish between them by using both subscripts and superscripts, which, however, do not represent covariant and contravariant indices.) Therefore, their decomposition is consistent only with the use of the alternative bond order formula eqn (8).†

In the next section we will briefly describe the formalism of effective atomic densities,[20] which will allow us to derive in a straightforward manner the most appropriate expressions for the decomposition of $\langle \hat{S}^2 \rangle$ in the framework of Hilbert space analysis.

3 Effective atomic density matrices formalism

The formalism of the atomic and diatomic effective densities is based on the *exact* decomposition of one- and two-electron densities into components that can be considered their one-center and one- and two-center contributions, respectively. These atomic and diatomic densities are identified with the contributions of each atom and pairs of atoms to the overall density, and can be used to derive in a common framework atomic populations, bond orders, atomic valences, molecular energy components, *etc*. for any kind of *atom in molecule* definition. In the simplest case, one can define the effective atomic contributions to the electron density, $\rho_A(\vec{r})$, simply fulfilling $\rho(\vec{r}) \equiv \sum_A^N \rho_A(\vec{r})$. The integral over the whole space of this function for atom A quite naturally yields the electron population associated with the atom

$$\int \rho_A(\vec{r}) \mathrm{d}\vec{r} \; = \; N_A. \tag{9}$$

The actual numerical value depends upon how $\rho_A(\vec{r})$ is defined. In the framework of 3D-space analysis, $\rho_A(\vec{r})$ can be written in general as

$$\rho_A(\vec{r}) \; = \; w_A(\vec{r})\rho(\vec{r}) \tag{10}$$

where w_A is a non-negative weight function defined for each atom and each point of the 3D space satisfying $\sum_A w_A(\vec{r}) = 1$. The actual definition of *atom in the molecule* ("fuzzy" or disjoint) is contained in the atomic weight functions.

In the case of Hilbert-space analysis, the effective atomic density can be most suitable written in terms of the matrix elements of the LCAO density matrix as

$$\rho_A(\vec{r}) = \sum_{\mu \in A} \sum_{\nu} D_{\mu\nu} \chi_\nu^*(\vec{r}) \chi_\mu(\vec{r}). \tag{11}$$

It is trivial to see that the integration of eqn (11) yields Mulliken's gross population of atom A, in accord with eqn (6).

In a similar manner, by combining the appropriate effective atomic contributions of the first-order density matrix to build effective diatomic exchange densities:

$$\rho_x^{AB}(\vec{r};\vec{r}') \; = \; \tfrac{1}{2}\left[\rho_A(\vec{r};\vec{r}')\rho_B(\vec{r}';\vec{r}) \; + \; \rho_B(\vec{r};\vec{r}')\rho_A(\vec{r}';\vec{r})\right] \tag{12}$$

† It has recently been discovered that this type of decomposition had also been used by some of us[7] in the decomposition of $\langle \hat{S}^2 \rangle$ from a formula different from eqn (1) as a result of a programming error: two subscripts have been interchanged by a mistake in the treatments of the cumulants.

where

$$\rho_A(\vec{r},\vec{r}') = \sum_{\mu \in A} \sum_{\nu} D_{\mu\nu} \chi_\nu^*(\vec{r}) \chi_\mu(\vec{r}') \tag{13}$$

is the atomic component of the first order spin-less density matrix $\rho(\vec{r},\vec{r}')$. One can easily recover upon integration of eqn (12) the expression of eqn (7) for the Mayer–Wiberg (closed-shell) bond order

$$\begin{aligned} \iint \rho_x^{AB}(\vec{r},\vec{r}')\mathrm{d}\vec{r}\,\mathrm{d}\vec{r}' \; &= \; \frac{1}{2}\sum_{\mu \in A}\sum_{\nu} D_{\mu\nu}S_{\nu\sigma}\sum_{\sigma \in B}\sum_{\lambda} D_{\sigma\lambda}S_{\lambda\mu} \\ &+ \frac{1}{2}\sum_{\mu \in B}\sum_{\nu} D_{\mu\nu}S_{\nu\sigma}\sum_{\sigma \in A}\sum_{\lambda} D_{\sigma\lambda}S_{\lambda\mu} \\ &= \sum_{\mu \in A}\sum_{\sigma \in B} (\boldsymbol{DS})_{\mu\sigma}(\boldsymbol{DS})_{\sigma\mu} = B_{AB} \end{aligned} \tag{14}$$

and so forth.

One advantage of using the formalism of effective atomic densities is that one can switch from 3D-space to Hilbert-space formulae or *vice versa* simply by taking the appropriate form of the effective densities involved in the calculation. In a recent paper[15] we have put forward an improved general formula for the decomposition of $\langle \hat{S}^2 \rangle$ applicable for both single-determinant and correlated wave functions. The numerical implementation of the resulting one- and two-center components was originally carried out in the 3D-physical space. Here we will make use of the formalism of the atomic and diatomic effective matrices depicted above to derive in a simple manner the appropriate one and two-center components of $\langle \hat{S}^2 \rangle$ in the framework of Hilbert-space analysis.

This exercise is of particular interest here because of the formal ambiguity affecting Hilbert-space decompositions in the selection of the indices put forward in the previous section. We most definitely recommend to stick to the assignment of subscripts that will be obtained here, which is consistent with both Mulliken population analysis and the original bond order definition eqn (7).

4 Decomposition of $\langle \hat{S}^2 \rangle$

Within the formalism of the effective atomic densities, the respective one- and two-center contributions to $\langle \hat{S}^2 \rangle$ can be formally written from the general expression of eqn (1) simply as

$$\begin{aligned} \langle \hat{S}^2 \rangle_A \; = \; &\tfrac{3}{4}\int u_A(\vec{r}_1)\mathrm{d}\vec{r}_1 \; + \; \tfrac{1}{2}\iint[\Gamma_{AA}(\vec{r}_1,\vec{r}_2) \\ &- \tfrac{1}{2}\rho_A^s(\vec{r}_1;\vec{r}_2)\rho_A^s(\vec{r}_2;\vec{r}_1)]\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 - \tfrac{1}{2}\iint[\Gamma_{AA}(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1) \\ &- \tfrac{1}{2}\rho_A^s(\vec{r}_1;\vec{r}_1)\rho_A^s(\vec{r}_2;\vec{r}_2)]\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 \end{aligned} \tag{15}$$

and

$$\begin{aligned} \langle \hat{S}^2 \rangle_{AB} \; = \; &\tfrac{1}{2}\iint[\Gamma_{AB}(\vec{r}_1,\vec{r}_2) - \tfrac{1}{2}\rho_A^s(\vec{r}_1;\vec{r}_2)\rho_B^s(\vec{r}_2;\vec{r}_1)\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2 \\ &- \tfrac{1}{2}\iint[\Gamma_{AB}(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1) - \tfrac{1}{2}\rho_A^s(\vec{r}_1;\vec{r}_1)\rho_B^s(\vec{r}_2;\vec{r}_2)]\,\mathrm{d}\vec{r}_1\mathrm{d}\vec{r}_2, \end{aligned} \tag{16}$$

where the atomic (in the case of $u(\vec{r})$ and $\rho^s(\vec{r}_1,\vec{r}')$) and diatomic (in the case of the cumulants, Γ) densities have been conveniently used, instead of their global counterparts in eqn (1).

In order to be consistent with the definitions of eqn (11) in the framework of Hilbert-space analysis, the effective atomic contributions of the density of effectively unpaired electrons and the spin density matrix must be taken as

$$\rho^{s}_{A}(\vec{r};\ \vec{r}') = \sum_{\mu\in A}\sum_{\nu} \boldsymbol{P}^{s}_{\mu\nu}\chi^{*}_{\nu}(\vec{r})\chi_{\mu}(\vec{r}') \tag{17}$$

and

$$u_{A}(\vec{r}) = \sum_{\mu\in A}\sum_{\nu} [2D_{\mu\nu} - (\boldsymbol{DSD})_{\mu\nu}]\chi^{*}_{\nu}(\vec{r})\chi_{\mu}(\vec{r}), \tag{18}$$

where we have made use of eqn (2).

The spin-less cumulant, Γ, being a genuine two-electron quantity, consists of atomic (if A = B) and diatomic (A ≠ B) contributions:

$$\Gamma_{AB}(\vec{r}_1,\vec{r}_2;\vec{r}'_1,\vec{r}'_2) = \sum_{\mu\in A}\sum_{\sigma\in B}\sum_{\nu,\lambda}\Gamma_{\mu\sigma\nu\lambda}\chi^{*}_{\nu}(\vec{r}_1)\chi^{*}_{\lambda}(\vec{r}_2)\chi_{\mu}(\vec{r}'_1)\chi_{\sigma}(\vec{r}'_2) \tag{19}$$

where $\Gamma_{\mu\sigma\nu\lambda}$ are the corresponding matrix elements of the cumulant in the atomic orbital basis.

Substituting eqn (17)–(19) into (15) and integrating, one obtains, after some manipulations, the final expression for the atomic components of $\langle\hat{S}^2\rangle$

$$\begin{aligned}\langle\hat{S}^2\rangle_{A} &= \frac{3}{4}\sum_{\mu\in A}[2(\boldsymbol{DS})_{\mu\mu} - (\boldsymbol{DSDS})_{\mu\mu}] - \frac{1}{4}\sum_{\mu,\nu\in A}(\boldsymbol{P}^{s}\boldsymbol{S})_{\mu\nu}(\boldsymbol{P}^{s}\boldsymbol{S})_{\nu\mu}\\ &+\frac{1}{4}\sum_{\mu,\nu\in A}(\boldsymbol{P}^{s}\boldsymbol{S})_{\mu\mu}(\boldsymbol{P}^{s}\boldsymbol{S})_{\nu\nu}\\ &+\frac{1}{2}\sum_{\mu,\sigma\in A}\sum_{\nu,\lambda}(\Gamma_{\mu\sigma\nu\lambda} - \Gamma_{\mu\sigma\lambda\nu})S_{\lambda\sigma}S_{\nu\mu}.\end{aligned} \tag{20}$$

Similarly, for the diatomic spin components one gets

$$\begin{aligned}\langle\hat{S}^2\rangle_{AB} &= -\frac{1}{4}\sum_{\mu\in A}\sum_{\nu\in B}(\boldsymbol{P}^{s}\boldsymbol{S})_{\mu\nu}(\boldsymbol{P}^{s}\boldsymbol{S})_{\nu\mu}\\ &+\frac{1}{4}\sum_{\mu\in A}\sum_{\nu\in B}(\boldsymbol{P}^{s}\boldsymbol{S})_{\mu\mu}(\boldsymbol{P}^{s}\boldsymbol{S})_{\nu\nu}\\ &+\frac{1}{2}\sum_{\mu\in A}\sum_{\sigma\in B}\sum_{\nu,\lambda}(\Gamma_{\mu\sigma\nu\lambda} - \Gamma_{\mu\sigma\lambda\nu})S_{\lambda\sigma}S_{\nu\mu}.\end{aligned} \tag{21}$$

In the single-determinant case the cumulants vanish and these formulae reduce to those derived independently in ref. 27.

Eqn (13) and (14) of Alcoba *et al.*[9] are similar to our eqn (20) and (21). In that paper, however, aside from the fact that the authors started from a formula different from eqn (1) – therefore the coefficients of the different terms of eqn (20) and (21) are different – the authors also chose a different convention in the treatment of the cumulant part: one index of the overlap matrix is assigned according to the term coming from the "bra" and another coming from the "ket". From now on we will refer to that different convention as formula with "interchanged" indices.

Hilbert-space analysis is not restricted to Mulliken's recipe. Another alternative scheme is Löwdin population analysis,[28] in which the atomic orbitals are first transformed to an orthogonal basis. Even though it is less often used, Löwdin analysis typically exhibits less basis set effects than Mulliken's. Indeed, it is well-known that Mulliken-based analyses can yield meaningless results if combined with diffuse functions lacking marked atomic character.[29] In the Löwdin basis the overlap matrix is a unit matrix and, as a consequence, the schemes with conventional and "interchanged" indices are equivalent.

5 Numerical results

We have written a program that performs the decomposition of $\langle\hat{S}^2\rangle$ described above in the framework of the Hilbert-space analysis for both Mulliken and Löwdin schemes, using the appropriate *effective* atomic overlap matrices outlined in the Appendix. Since Löwdin analysis is not strictly rotational invariant[30] with Cartesian 6d atomic orbitals, we recommend its use only with pure 5d orbitals. We have included results with 6d functions only for comparison purposes. The first- and second-order density matrices have been obtained using a modified version of Gaussian-03 program suite[31] and an auxiliary program[32] that reads and processes CISD and CASSCF outputs. All calculations have been carried out with the geometrical structure of the molecules optimized at the current level of theory unless otherwise stated.

To assess the numerical effect on the use of the different summation schemes in the Hilbert-space analysis, we have studied the H_2 molecule at the CASSCF(2,4) level for several basis sets. Table 1 gathers the local spin values on the H atom for the conventional $\langle\hat{S}^2\rangle_{H}$ and "interchanged" index conventions $\langle\hat{S}^2\rangle^{int}_{H}$ within Mulliken's scheme. The values for Löwdin $\langle\hat{S}^2\rangle^{L}_{H}$ and 3D-space analysis $\langle\hat{S}^2\rangle^{3D}_{H}$ (using Becke atoms) are also included for comparison.

The local spin values using eqn (20) are close to zero in all cases, in line with the physical expectations. The numbers exhibit reasonably small basis set dependence and are also very similar to those obtained in the framework of 3D-space analysis. Using the alternative formula with "interchanged" indices the values are somewhat too large, as compared with the "conventional" ones, and suffer from strong basis set effects, especially when combining two sets of diffuse functions and Cartesian 6d 10f orbitals. With this extended basis set

Table 1 Atomic local spin values calculated at the CASSCF(2,4) level for the H_2 molecule at interatomic distance R_{H-H} = 0.746 Å for several basis sets

Basis set	$\langle\hat{S}^2\rangle_{H}$	$\langle\hat{S}^2\rangle^{int}_{H}$	$\langle\hat{S}^2\rangle^{L}_{H}$	$\langle\hat{S}^2\rangle^{3D}_{H}$
cc-pVDZ	0.036	0.060	0.028	0.034
cc-pVTZ	0.035	0.051	0.025	0.034
cc-pVQZ	0.039	0.066	0.027	0.036
cc-pVTZ (6d)	0.035	0.052	0.027	0.034
cc-pVQZ (6d 10f)	0.043	0.089	0.028	0.036
aug-cc-pVDZ	0.038	0.069	0.025	0.034
aug-cc-pVTZ	0.020	0.031	0.026	0.036
aug-cc-pVQZ	0.035	0.044	0.026	0.036
d-aug-cc-pVQZ	0.045	0.130	0.024	0.034
aug-cc-pVTZ (6d)	0.046	0.115	0.027	0.036
aug-cc-pVQZ (6d 10f)	0.055	0.192	0.026	0.036
d-aug-cc-pVQZ (6d 10f)	0.095	0.891	0.023	0.034

even the $\langle \hat{S}^2 \rangle_H$ value (0.095) is considerably larger than the rest. The $\langle \hat{S}^2 \rangle_H^L$ and $\langle \hat{S}^2 \rangle_H^{3D}$ values show virtually no basis set dependency. However, it seems that for the smaller basis sets Mulliken's values are closer to the 3D-space ones than Löwdin's.

The recommended Mulliken-type decomposition has also been applied to a series of singlet molecules and the results are presented in Table 2. The optimized geometries and the wavefunctions were obtained at the CISD/6-31G** level of theory (with Cartesian 6d functions). Note that for these systems the overall $\langle \hat{S}^2 \rangle$ value is zero, but small local atomic spins can be induced by correlation fluctuations. One should only expect the presence of significant diatomic contributions in singlet systems if there would be any antiferromagnetic coupling that could be distinguished from covalent bonding.

As anticipated, the molecules of the H_nX series (HF, H_2O, and NH_3) show small values of local spin. The only systems with atomic spin contributions larger than 0.1 are homonuclear diatomic Li_2 and Be_2, with $\langle \hat{S}^2 \rangle_{Li} = 0.156$ and $\langle \hat{S}^2 \rangle_{Be} = 0.175$, respectively. These values are consistent with those obtained within the framework of 3D-space analysis.[15] On the other hand, in the series of hydrocarbons the local spins on the C atoms reported here show relevant differences. Within the 3D-space formulation the atomic spin contributions were always below 0.1.[15] In the Hilbert-space framework both CH_4 ($\langle \hat{S}^2 \rangle_C = 0.320$) and C_2H_6 ($\langle \hat{S}^2 \rangle_C = 0.199$) present quite significant local spin on the C atoms. The most striking finding in Table 2 is the negative local spin on the C atom obtained for C_2H_2. Since there is no physical explanation for a negative value of $\langle \hat{S}^2 \rangle_C$, we tried to understand this odd behavior. The local spin obtained at the same level of theory with the 3D-space formulation was 0.083,[15] which rules out the truncated CISD wave function as responsible for the spurious number. In order to check for basis set effects on the local spin for this system, we have computed the $\langle \hat{S}^2 \rangle_C$ values at the CISD/6-31G** optimized geometry using several basis sets.

Table 2 CISD/6-31G** atomic $\langle \hat{S}^2 \rangle_A$ and diatomic $\langle \hat{S}^2 \rangle_{AB}$ values for a set of singlet molecules at optimized geometries

Molecule		$\langle \hat{S}^2 \rangle_A / \langle \hat{S}^2 \rangle_{AB}$	Molecule		$\langle \hat{S}^2 \rangle_A / \langle \hat{S}^2 \rangle_{AB}$
H_2	H	0.036	C_2H_6	C	0.199
	H–H	−0.036		H	0.024
Li_2	Li	0.156		C–C	−0.122
	Li–Li	−0.156		C–H	−0.069
Be_2	Be	0.175		C···H	0.034
	Be–Be	−0.175		H–H	0.018
HF	H	0.006		H···H	−0.015
	F	0.006	C_2H_4	C	0.056
	H–F	−0.006		H	0.024
H_2O	H	0.013		C–C	−0.094
	O	0.013		C–H	−0.036
	O–H	−0.007		C···H	0.055
	H···H	−0.006		H–H	−0.002
NH_3	N	0.061		$H\cdots H_{cis}$	−0.026
	H	0.019		$H\cdots H_{trans}$	−0.014
	N–H	−0.020	C_2H_2	C	*−0.139*
	H···H	0.000		H	0.019
CH_4	C	0.320		C–C	*0.147*
	H	0.026		C–H	0.048
	C–H	−0.080		C···H	−0.056
	H···H	0.018		H···H	−0.006

Table 3 Atomic $\langle \hat{S}^2 \rangle_C$ components for acetylene molecule computed at the CISD level of theory with different basis sets

Basis set	$\langle \hat{S}^2 \rangle_C$	$\langle \hat{S}^2 \rangle_C^L$	$\langle \hat{S}^2 \rangle_C^{3D}$
STO-3G	0.012	0.179	0.159
6-31G	−0.074	0.119	0.114
6-31G**	−0.136	0.082	0.084
6-31G**(6d)	−0.139	0.080	0.083
6-311G	0.030	0.119	0.113
6-311G**	−0.118	0.085	0.085
6-311G**(6d 10f)	−0.143	0.091	0.084
cc-pVDZ	−0.162	0.078	0.085
cc-pVTZ	−0.271	0.087	0.079
cc-pVTZ(6d 10f)	−0.146	0.104	0.078
aug-cc-pVDZ	0.729	0.078	0.087
aug-cc-pVTZ	−2.926	0.092	0.078

The results are gathered in Table 3. The results using Löwdin and 3D-space schemes are also included. The $\langle \hat{S}^2 \rangle_C$ is still negative for most basis sets, except for the STO-3G, 6-311G and aug-cc-pVDZ basis sets. Note the completely meaningless value of −2.926 obtained with Mulliken's scheme with the quite standard aug-cc-pVTZ basis set.‡ Moreover, since negative values appear for both small and relatively large basis sets with and without polarization functions it is difficult to draw any general conclusion from the data. However, the local spin on C atoms using both Löwdin's and the 3D-space formulation is always small and positive, as it should be. One can see a systematic lowering of the value upon inclusion of polarization functions and no significant effects of the diffuse functions.

In this context it is worth noting that rather odd numbers have been obtained in the literature when combining Mulliken analysis techniques and genuine two-electron quantities from correlated wave functions. For instance, Vyboishchikov *et al.*[33] found unphysically positive correlation contributions for diatomic energies at the CID level of theory. On the other hand, the bond order indices that in the correlated case make use of the actual pair density (the so-called delocalization index, DI[34–39]) have also been a matter of debate as, for the simplest case of H_2 described with Weinbaum's classical correlated wave function, the DI gives just 0.39.[38] One can conclude once again that in some cases one can get spurious results when decomposing quantities that explicitly include the second-order density matrix in the framework of Mulliken analysis. Fortunately, according to our experience, the pathological case of acetylene seems to be quite exceptional (*e.g.* no such problems occur for the isoelectronic N_2 molecule). Nevertheless, the Hilbert-space results using Löwdin's scheme prove to be much more robust and reliable, especially for large basis sets.

Finally, we have also considered several radical (doublet) and diradical singlet molecules. In principle, the magnitude of the local spin values and the diatomic spin components, compared to the ideal values for localized spins, can help to quantify the diradical character of the molecule. For a system of two perfectly localized anti-parallel spins on centers A and B,

‡ For this peculiar molecule, a non-nuclear attractor is found for the STO-3G and cc-pVTZ basis sets, having no apparent effect on the local spin values.

Table 4 Atomic $\langle \hat{S}^2 \rangle_C$ and diatomic $\langle \hat{S}^2 \rangle_{CC}$ components at the CASSCF/6-31G** level of theory. Active spaces used are (6,6) for benzene, (7,7) for phenyl radical, (8,8) for *ortho*-, *meta*- and *para*-benzyne and (3,3) for the allyl radical

	C_1	C_2	C_3	C_4	C_5	C_6
Benzene						
C_1	0.114	−0.101	0.079	−0.069		
Phenyl radical						
C_1	0.968	−0.145	0.100	−0.103		
C_2		0.129	−0.102	0.079	−0.071	0.078
C_3			0.127	−0.102	0.080	
C_4				0.118		
o-Benzyne						
C_1	0.324	−0.304	0.086	−0.075	0.080	−0.110
C_3			0.130	−0.107	0.079	
C_4				0.115	−0.092	
m-Benzyne						
C_1	0.540	−0.101	−0.344	−0.056	0.069	−0.104
C_2		0.133		0.066	−0.065	
C_4				0.116	−0.097	0.074
C_5					0.121	
p-Benzyne						
C_1	0.962	−0.171	0.125	−0.862		
C_2		0.145	−0.106		−0.072	0.078
Allyl radical						
C_1	0.440	−0.145	0.151			
C_2		0.145				

a proper (without spin contamination) wave function would yield local values of $\langle \hat{S}^2 \rangle_A = \langle \hat{S}^2 \rangle_B = 3/4$ and $\langle \hat{S}^2 \rangle_{AB} = -3/4$, for an overall value of $\langle \hat{S}^2 \rangle = 0$, characteristic of a singlet. In the case of a molecular system with a singly localized unpaired electron on a center A, one can expect a local value of $\langle \hat{S}^2 \rangle_A$ close to 3/4 (if this ideal system is described at the ROHF level of theory then $\langle \hat{S}^2 \rangle_A \equiv 3/4$ and all diatomic spin contributions exactly vanish).[15]

The local spin analysis for benzyl and allyl radicals and the set of diradical *ortho*-, *meta*- and *para*-benzyne molecules in the singlet electronic state are gathered in Table 4. Benzene molecule has also been included for comparison. The systems have been studied at the CASSCF/6-31G** level of theory with appropriate active spaces, *i.e.*, full π-valence and appropriate σ orbitals for the radical systems. We report only the results obtained for Mulliken-type analysis. Using Löwdin or 3D-space schemes does not change significantly the values for these systems.

Non-negligible local spin (0.114) is observed on the C atoms of benzene, due to the fluctuation of π-electrons induced by the electron correlation. The sign of the diatomic contributions is consistent with the chemical picture displayed in Fig. 1a. Such localization of the spins is also consistent with the observed decrease in the electronic aromaticity indices upon inclusion of electron correlation[40] (note that a single-determinant restricted description of benzene gives identically zero local spins). Of course, the magnitude of the local spin and the diatomic spin contributions is very small compared to the ideal 3/4 value, as one could anticipate for a genuine diamagnetic system.

In phenyl radical the unpaired electron is localized on a σ orbital at/near C_1 (in which the H atom is absent). The local spin in the remaining C atoms is slightly greater than that in the case of benzene, as well as the magnitude and sign pattern of the diatomic spin components. The local spin value of the

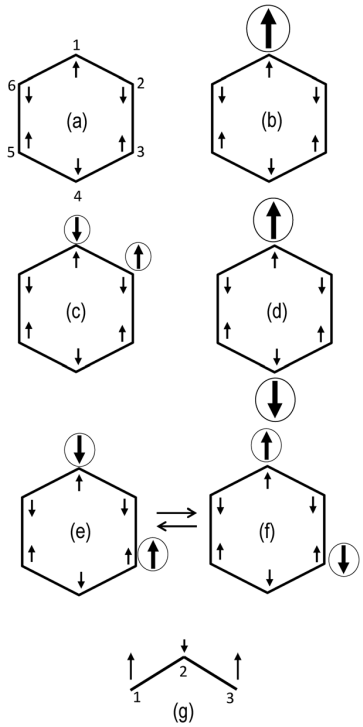

Fig. 1 Localized spins picture emerging from the local spin analysis. Circled arrows represent the σ contributions, small arrows indicate the π counterpart.

radical center (0.968) exceeds the *ideal* value of a singly localized electron (3/4). Furthermore, the $\langle \hat{S}^2 \rangle_{12}$ value increases (in absolute value), as well as the local spin on C_2 with respect to the benzene. Of course, one cannot expect these contributions to be fully additive as the shape of the π-orbitals in the phenyl radical system is modified by the presence of the unpaired σ electron, and there can also be some σ–π interplay. Nevertheless, one can conclude that there must be a parallel alignment of the local spin arising from the π and σ electrons in the radical center, as indicated in Fig. 1b.

The analysis of the allyl radical gives a completely different picture. In this case there are significant local spin contributions from all C atoms. The main local spin centers are C_1 and C_3 atoms, with a value of 0.440, and the central C_2 atom also contributes 0.145. The diatomic terms reveal partial anti-ferromagnetic coupling (−0.145) between C_1 and C_2, and a ferromagnetic one of similar magnitude (0.151) between C_1 and C_3. It is worth noting that the diatomic spin components would vanish for a ROHF description of this system. Thus, it is clear that in this case there is a significant interplay between the three π electrons of the system, as depicted in Fig. 1g.

Among the set of diradical benzyne isomers, *para*-benzyne exhibits local spin features similar to the phenyl radical. The local spin on the radical centers (atoms C_1 and C_4) is 0.962 and the diatomic spin contribution $\langle \hat{S}^2 \rangle_{14} = -0.862$ indicates anti-ferromagnetic coupling. The $\langle \hat{S}^2 \rangle_{12}$ value is similar to that obtained for the phenyl radical, and the diatomic values involving the remaining atoms with contributions from the

π orbitals are close to the case of benzene. The overall picture of this system is consistent with a σ,σ-diradical, as depicted in Fig. 1d.

The diradical character decreases in the case of *ortho*- and *meta*-benzyne. In the former, the local spin on atoms C_1 and C_2 is just 0.324, substantially smaller than the ideal value of 3/4 for a fully localized electron. Concomitantly, the diatomic $\langle \hat{S}^2 \rangle_{12}$ value decreases to −0.304. Since both values practically compensate each other, the local spin contribution from the σ electrons is largely confined in these two atoms. Therefore, the small magnitude of the local spin indicates that the two electrons must exhibit some genuine pairing, leading to an increase in the bonding interaction between the two atoms relative to benzene. Indeed, the computed bond orders for *ortho*-benzyne and benzene are 2.12 and 1.38, respectively. The diatomic terms involving the remaining atoms are again quite similar to the case of benzene. The negative $\langle \hat{S}^2 \rangle_{16}$ value of −0.110 could have its origin merely in the π electron system. Hence, it is difficult in this case to establish whether the picture of the system is the one displayed in Fig. 1c or the one with the σ spins of atoms C_1 and C_2 interchanged. We have also performed the analysis for a CASSCF(2,2) wave function (at the same geometry) where only the σ electrons are correlated. In this case the local spin on the C_1 atom is just 0.271. The estimate of 0.11–0.13 contribution of the π part for this system seems to point towards an anti-parallel arrangement of the σ and π electrons on the C_1 and C_2 atoms, as depicted in the figure.

The *meta*-benzyne molecule represents an intermediate situation between the *ortho* and *para* isomers. There are however some interesting features. First of all, the local spin on the C_1 and C_3 atoms amounts to an intermediate value of 0.540, but the diatomic term is just −0.304. Intuitively one could conclude that the contribution of the σ to the local spin is not fully compensated between the two atoms, as in the case of *ortho*-benzyne. However, since the contribution of the π electrons to the $\langle \hat{S}^2 \rangle_{13}$ value is *positive* (indicating parallel alignment), it partially compensates for a larger negative contribution from the anti-parallel arrangement of the σ electrons. In this sense it is noticeable that for this molecule the $\langle \hat{S}^2 \rangle_{12}$ value (−0.101) is exactly the same as in benzene, and also very similar to the $\langle \hat{S}^2 \rangle_{16}$ one (−0.104). However, due to symmetry, a parallel or anti-parallel alignment of the σ and π local spin contributions is not possible for both atoms C_1 and C_3, as shown in Fig. 1e and f. As these atoms are equivalent by symmetry, the picture of the localized spins of this system must be a combination of the two (equivalent) configurations. This suggests that the σ and π contributions to the local spin of this molecule could be additive to a large extent. Indeed, the values of the atomic and diatomic terms involving C_2, C_4, C_5 and C_6 centers are very similar to those of benzene. Also, the local spin on the C_1 atom for a CASSCF(2,2) wave function is 0.469. After adding the estimate contribution of the π part (*ca.* 0.1) one gets a value that is only slightly larger than the actual value of 0.540.

In summary, the diradical character of the three isomers increases from *ortho* to *para*, in agreement with other analysis.[41] The local spin analysis allows for a deeper insight into how the local spins are distributed in the centers and its magnitude.

6 Conclusions

We have carried out the decomposition of the expectation value of the spin operator for a general wave function in the framework of the Hilbert-space analysis that fulfills all the requirements imposed to date. We have shown that there is an ambiguity affecting Mulliken-type decompositions in the selection of the indices where the atoms are centered. We definitely recommend to stick to the assignment of subscripts that is consistent with Mulliken population analysis and the original Mayer–Wiberg bond orders. The results obtained are in good agreement with physical expectation and, in general, do not depend too much on the basis set. For the particular case of acetylene one can get spurious results when performing the decomposition of quantities that explicitly depend upon the second-order density matrix, even with small basis sets. We show that Löwdin's scheme is more robust and reliable in all cases.

7 Appendix

7.1 Mapping between 3D and Hilbert space analyses

One can introduce a mapping between the atomic overlap matrices (used in the framework of 3D-space analysis) and their Hilbert space analogues in order to establish a one-to-one correspondence between the expressions obtained for 3D-space and Hilbert-space analyses.

Let us consider for simplicity a closed-shell system with doubly occupied molecular orbitals. In the framework of 3D-space analysis, the gross atomic population of atom A is obtained as

$$N_A = \int w_A(\vec{r})\rho(\vec{r})d\vec{r} = \sum_{\mu\nu} D_{\mu\nu}S^A_{\nu\mu} = 2\sum_{\mu\nu}\sum_{i}^{occ} c^*_{\nu i}S^A_{\nu\mu}c_{\mu i} = 2\,\mathrm{tr}(\boldsymbol{C}^+\boldsymbol{S}^A\boldsymbol{C}), \quad (22)$$

where

$$S^A_{\nu\mu} = \int w_A(\vec{r})\chi^*_\nu(\vec{r})\chi_\mu(\vec{r})d\vec{r} \quad (23)$$

are the elements of the atomic overlap matrix $\boldsymbol{S}^A$ and $\boldsymbol{C}$ is the matrix containing the orbital coefficients of the occupied molecular orbitals. The Mulliken gross population defined in eqn (6) can be also written in terms of the molecular orbital coefficients as

$$N_A = 2\sum_{\mu\in A}\sum_{\nu}\sum_{i}^{occ} c^*_{\nu i}S_{\nu\mu}c_{\mu i} = 2\,\mathrm{tr}(\boldsymbol{C}^+\boldsymbol{S}\eta^A\boldsymbol{C}) \quad (24)$$

where η^A is a block-truncated unit matrix with all elements equal to zero except $\eta^A_{\mu\mu} = 1$ for $\mu \in A$. Comparing eqn (22) and (24) one can put into correspondence with the atomic overlap matrix in atomic orbital basis $\boldsymbol{S}^A$ the matrix product $\boldsymbol{S}\eta^A$. It is easy to see that if one expresses one- and two-center terms of ref. 15 in the atomic orbital basis and replaces the matrix elements $S^A_{\mu\nu}$ by the $[\boldsymbol{S}\eta^A]_{\mu\nu}$ ones, the one- and two-center terms of eqn (20) and (21) can be recovered. Note that this Mulliken's *effective* atomic overlap matrix is non-symmetric, and this is the reason why the "interchanged" indices convention

mentioned above leads to different expressions for the one- and two-center components. In that scheme one essentially uses both $\boldsymbol{S}\eta^{\mathrm{A}}$ and $\eta^{\mathrm{A}}\boldsymbol{S}$ matrices. Thus, the one- and two-center contributions of $\langle\hat{S}^2\rangle$ in the framework of Löwdin analysis can be easily derived if an analogous mapping could be established. In Löwdin population analysis one has

$$N_{\mathrm{A}}^{\mathrm{L}} = \sum_{\mu\in\mathrm{A}}\sum_{\nu} S_{\nu\mu}^{1/2} D_{\mu\sigma} S_{\sigma\nu}^{1/2}, \tag{25}$$

which can also be expressed in terms of the MO coefficients as

$$N_{\mathrm{A}}^{\mathrm{L}} = 2\sum_{\mu\in\mathrm{A}}\sum_{\nu\sigma}\sum_{i}^{\mathrm{occ}} c_{\sigma i}^{*} S_{\sigma\mu}^{1/2} S_{\mu\nu}^{1/2} c_{\nu i} = 2\,\mathrm{tr}(\boldsymbol{C}^{+}\boldsymbol{S}^{1/2}\eta^{\mathrm{A}}\boldsymbol{S}^{1/2}\boldsymbol{C}) \tag{26}$$

Comparing this expression with eqn (22) and (24) it is easy to identify $\boldsymbol{S}^{1/2}\eta^{\mathrm{A}}\boldsymbol{S}^{1/2}$ as the appropriate Löwdin's *effective* atomic overlap matrix. Note that in this case, the atomic overlap matrix is symmetric, which means that in the framework of Löwdin analysis the conventional and ''interchanged'' index schemes are equivalent.

Acknowledgements

Financial help has been furnished by the Spanish MICINN Projects No. CTQ2011-23441/BQU, CTQ2011-23156/BQU and Acción Complementaria del MCI (PCI2006-A7-0631). Financial support from MICINN and the FEDER fund (European Fund for Regional Development) was also provided by grant UNGI08-4E-003. E.R.-C. acknowledges support from the Spanish FPU program (Grant No. AP2008-01231). E.M. acknowledges financial support from the EU under a Marie Curie Career Integration grant (PCI09-GA-2011-294240) and the Beatriu de Pinós program from AGAUR for the postdoctoral grant (BP_B_00236). I.M. acknowledges partial financial support from the Hungarian Scientific Research Fund (grant OTKA 71816).

References

1 A. Clark and E. Davidson, *J. Chem. Phys.*, 2001, **115**, 7382–7392.
2 A. Clark and E. Davidson, *Mol. Phys.*, 2002, **100**, 373–383.
3 A. Clark and E. Davidson, *J. Phys. Chem. A*, 2002, **106**, 6890–6896.
4 I. Mayer, *Chem. Phys. Lett.*, 2007, **440**, 357–359.
5 D. Alcoba, L. Lain, A. Torre and R. Bochicchio, *Chem. Phys. Lett.*, 2009, **470**, 136–139.
6 I. Mayer, *Chem. Phys. Lett.*, 2009, **478**, 323–326.
7 I. Mayer and E. Matito, *Phys. Chem. Chem. Phys.*, 2010, **12**, 11308–11314.
8 D. Alcoba, A. Torre, L. Lain and R. Bochicchio, *Chem. Phys. Lett.*, 2011, **504**, 11308–11314.
9 D. R. Alcoba, A. Torre, L. Lain and R. C. Bochicchio, *J. Chem. Theory Comput.*, 2011, **7**, 3560–3566.
10 M. Reiher, *Faraday Discuss.*, 2006, **135**, 97–124.
11 C. Herrmann, M. Reiher and B. Hess, *J. Chem. Phys.*, 2005, **122**, 034102.
12 A. V. Luzanov and O. V. Prezhdo, *Mol. Phys.*, 2007, **105**, 2879–2891.
13 M. Podewitz, C. Herrmann, A. Malassa, M. Westerhausen and M. Reiher, *Chem. Phys. Lett.*, 2008, **451**, 301–308.
14 C. Herrmann, L. Yu and M. Reiher, *J. Comput. Chem.*, 2006, **27**, 1223–1239.
15 E. Ramos-Cordoba, E. Matito, I. Mayer and P. Salvador, *J. Chem. Theory Comput.*, 2012, **8**, 1270–1279.
16 K. Takatsuka, T. Fueno and K. Yamaguchi, *Theor. Chim. Acta*, 1978, **48**, 175–183.
17 I. Mayer and P. Salvador, *Chem. Phys. Lett.*, 2004, **383**, 368–375.
18 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford Univ. Press, Oxford, 1990.
19 I. Mayer, *J. Comput. Chem.*, 2007, **28**, 204–221.
20 S. F. Vyboishchikov, P. Salvador and M. Duran, *J. Chem. Phys.*, 2005, **122**, 244110–244123.
21 R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1833–1841.
22 I. Mayer, *THEOCHEM*, 1992, **255**, 1–7.
23 I. Mayer, *Chem. Phys. Lett.*, 1983, **97**, 270–274.
24 H. Sato and S. Sakaki, *J. Phys. Chem. B*, 2007, **111**, 672.
25 R. Ponec and D. L. Cooper, *Int. J. Quantum Chem.*, 2004, **97**, 1002–1011.
26 J. G. Angyan, M. Loos and I. Mayer, *J. Phys. Chem.*, 1994, **98**, 5244–5248.
27 I. Mayer, *Chem. Phys. Lett.*, 2012, **539–540**, 172–174.
28 P.-O. Löwdin, *Adv. Quantum Chem.*, 1970, **5**, 185.
29 J. Baker, *Theor. Chim. Acta*, 1985, **68**, 221.
30 I. Mayer, *Chem. Phys. Lett.*, 2004, **393**, 209.
31 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03, Revision C.02*, Gaussian, Inc., Pittsburgh, PA, 2003.
32 E. Matito and F. Feixas, *DMN program*, University of Girona (Spain) and University of Szczecin, Poland, 2009.
33 S. Vyboishchikov and P. Salvador, *Chem. Phys. Lett.*, 2006, **430**, 204–209.
34 R. Bader and M. Stephens, *Chem. Phys. Lett.*, 1974, **26**, 445–449.
35 R. Bader and M. Stephens, *J. Am. Chem. Soc.*, 1975, **97**, 7391–7399.
36 J. Ángyán, E. Rosta and P. Surján, *Chem. Phys. Lett.*, 1999, **299**, 1–8.
37 M. Giambiagi, M. Giambiagi and F. Jorge, *Theor. Chim. Acta*, 1985, **68**, 337–341.
38 I. Mayer, *Int. J. Quantum Chem.*, 1986, **29**, 477–483.
39 X. Fradera, M. Austen and R. Bader, *J. Phys. Chem. A*, 1999, **103**, 304–314.
40 E. Matito, M. Solà, P. Salvador and M. Duran, *Faraday Discuss.*, 2007, **135**, 325–345.
41 A. Clark and E. Davidson, *J. Am. Chem. Soc.*, 2001, **123**, 10691–10698.

Chapter 5

The local spin analysis: Chemical bonding and radical character

5.1 Local spin and chemical bonding

Ramos-Cordoba, E.; Salvador, P.; Reiher, M. “Local Spin and Chemical Bonding”. *Chemistry A European Journal*. 2013, 19, 15267–15275

http://dx.doi.org/10.1002/chem.201300945

http://onlinelibrary.wiley.com/doi/10.1002/chem.201300945/abstract

Article first published online: 17 SEP 2013

Abstract

The electronic structure of main-group diatomic molecules is discussed in the light of local spin analysis. A deep investigation into the origin of local spins and their coupling is presented. It is shown that the presence of significant local spins in bonded molecules flags deviations from the classical bonding prototypes. For the notorious example of the C2 molecule, the local spin analysis indicates that its ground state has all ingredients to be categorized as a diradical.

Keywords

ab initio calculations; bond theory; electronic structure; local spin analysis; main group elements

5.2 Quantification of diradical character by the local spin analysis

As submitted for publication to Phys. Chem. Chem. Phys.

Diradical character from the local spin analysis

Eloy Ramos-Cordoba, Pedro Salvador
Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química,
University of Girona, 17071 Girona, Spain
e-mail: Pedro.Salvador@udg.edu

Abstract

Diradical species are analyzed on the light of the local spin analysis. The atomic and diatomic contributions to the overall $\langle \hat{S}^2 \rangle$ value are used to detect the diradical character of a number of molecular species mostly in their singlet state, for which no spin density exists. A general procedure for the quantification of diradical character for both singlet and triplet states is achieved by using a recently introduced index that measures the deviation of an actual molecule from an ideal system of perfectly localized spin centers. The index is of general applicability and can be easily determined in equal footing from a multireference or an open-shell single-determinant wave function.

Keywords: Local spin analysis, diradical character, benzyne isomers, propellane

Introduction

Salem[1] defined diradicals as molecules with two electrons occupying two near degenerate orbitals. Indeed, how close to degeneracy these orbitals are (HOMO-LUMO gap) or more generally the singlet-triplet gap is one of the characteristic features of diradical systems. Diradicals are important in chemistry since they emerge as intermediates of many chemical reactions.[2] Pure, ideal diradicals such as a dissociated H_2 singlet can be easily characterized theoretically from different indicators, depending on the nature of the wave function. However, the quantification of the diradical or diradicaloid[3] character of short-lived singlet diradicals is not so trivial because the formally unpaired electrons do interact to some extent. There is a continuum between the closed-shell spin-paired and the perfectly localized spin-entangled situations, as exemplified by the dissociation curve of singlet H_2.

Several indices have been proposed in the literature in order to detect and quantify the diradical character of molecular systems, the simplest probably being the value of $\langle \hat{S}^2 \rangle$ of a broken symmetry spin-unrestricted wave function.[4,5] For a system with an equal mixture of singlet and triplet components one should expect a $\langle \hat{S}^2 \rangle$ value close to 1.[6] Accordingly, Bachler et al.[7] proposed the following index

$$n_{rad} = 1 - \sqrt{1 - \langle \hat{S}^2 \rangle_{BS}}, \tag{1}$$

where $\langle \hat{S}^2 \rangle_{BS}$ represents a UHF broken-symmetry wave function. An alternative index can be built making explicit use of the occupation numbers of spin-unrestricted natural orbitals (UNOs). In a system with diradical character, a pair of bonding an antibonding orbitals are typically associated with the two radical sites. The closer to 1 the occupation of the antibonding orbital is, the higher the diradical character. Jung and Head-Gordon[3] and Bachler *et al.* [8] used the occupation numbers obtained from perfect-pairing approaches and Lopez et al. used the occupation number computed at the natural orbital functional (NOF) level of theory[9] to assess the extent of diradical character of different molecules. Rivero *et al.* also studied the extent of radical character from the occupation numbers that are close to one from a spin-

projected Hartree-Fock calculation.[10] Similarly, Kamada et al.[11] used the index,

$$y = \frac{(1-T)^2}{1+T^2} \quad \text{and} \quad T = \frac{n_{HOMO} - n_{LUMO}}{2} \tag{2}$$

where n_{HOMO} and n_{LUMO} are the occupations of the bonding and antibonding UNOs. In a purely closed-shell system $n_{HOMO} = 2$ and $n_{LUMO} = 0$, and hence $y = 0$. When the occupations of the two orbitals are equal the system is a pure diradical and $y = 1$.

When a multiconfigurational wave function is used the occupation numbers of the orbitals of the radical sites can be replaced by the weights of appropriate configurations of the CI expansion. In the simplest two-electrons in two-orbitals (the so-called magnetic orbitals) model, the 2x2 CI wave function is build up from a configuration in which the bonding combination of the magnetic orbitals is doubly occupied, and another that includes the double excitation to the antibonding combination of the magnetic orbitals. Bachler et. al.[7] proposed the following indicator for diradical character

$$n_{rad}^{CI} = \sqrt{2}\,|c_d| \tag{3}$$

where c_d is the weight of the doubly-excited configuration. Later on, other authors suggested an improved version[12] that also incorporates the weight of the other configuration

$$d = 2\sqrt{\frac{c_0^2 c_d^2}{c_0^2 + c_d^2}}. \tag{4}$$

None of the indices described above is of general applicability. Beyond diradicals a signature of polyradical character may be derived from the shape and occupation of the natural orbitals.[10,13] The applicability of the indices given in Eqns. (1) to (4) is thus restricted to diradical systems that can be well described with a two-electron two-orbital model. A noteworthy alternative is the analysis of the so-called density of effectively unpaired electrons, $u(\vec{r})$, defined by Takatsuka *et al.* [14] as

$$u(\vec{r}) = 2\rho(\vec{r}) - \int \rho(\vec{r};\vec{r}\,')\rho(\vec{r}\,';\vec{r})d\vec{r}\,'. \tag{5}$$

This quantity can be easily obtained at any level of theory from the first-order density matrix, $\rho(\vec{r};\vec{r}\,')$, and provides a spatial distribution of the unpaired or "odd"

electrons in the system, even if the spin density vanishes (*e. g.* for multiconfigurational singlet wave functions). The total number of unpaired electrons, N_D, can be recovered upon integration of $u(\vec{r})$ over the whole space. The topology $u(\vec{r})$ and the N_D values have been used by Staroverov and Davidson to analyze the evolution of the radical character upon a chemical reaction, *e.g.* the Cope rearrangement.[15,16] Cheng and Hu[17] found a good correlation between N_D and the singlet-triplet gap for a set of B_2P_2 ring derivative diradicaloids. Moreover, population analysis techniques such as Mulliken[16] or QTAIM[18] have also been applied to recover the average number of unpaired electrons on a given atom/fragment. It is worth to note that Mulliken populations of $u(\vec{r})$ are identical to Mayer's free valence index[19,20] for singlet wave functions.

In singlet diradicals the presence of some local spin associated to a given atom or fragment of the molecular system is assumed. The spin properties of molecular systems are usually characterized by the analysis of the spin density. In fact, spin-unrestricted single-determinant calculations often result in broken-symmetry solutions with non-vanishing spin density. In this case, however , the state of the system is not described as a pure singlet, as it appears contaminated with higher spin states. When a proper multireference wave function is used to describe a pure singlet the spin density exactly vanishes at all points of the space. Yet, one can still invoke the concept of local spin in the system.

Local spins can be retrieved from wave function analysis by a number of decomposition schemes.[21–27] The most appropriate approach to the problem, as pointed out by Mayer,[23] is probably the exact decomposition of the expectation value of the spin-squared operator into a sum of atomic and diatomic contributions as

$$\langle \widehat{S^2} \rangle = \sum_A \langle \widehat{S^2} \rangle_A + \sum_{A,B \neq A} \langle \widehat{S^2} \rangle_{AB}. \tag{6}$$

A proper formulation of eqn. (6) can provide vanishing one- and two-center terms for restricted single-determinant wave functions (thereby distinguishing electron pairing in bonds from antiferromagnetic coupling), and non-zero ones for pure singlets described by correlated wave functions, thus overcoming the limitation of use of the spin density. The actual expressions for the one- and two-center contributions fulfilling these conditions, henceforth local spin analysis, can be found elsewhere.[27]

In the local spin analysis, the $\langle \hat{S}^2 \rangle_A$ values indicate and quantify the presence of local spin *within* the molecule, namely on atom/fragment A. The magnitude and sign of the diatomic contributions $\langle \hat{S}^2 \rangle_{AB}$ with $B \neq A$ inform about the nature of the couplings between these local spins.[21,28,29] The physical interpretation of the $\langle \hat{S}^2 \rangle_A$ and particularly $\langle \hat{S}^2 \rangle_{AB}$ values is somewhat intricate, and has been recently discussed in detail in several papers.[30,31]

The ability of both local spin methods and the density of effectively unpaired electrons to capture the diradical nature of molecular systems has already been discussed in the recent literature.[15,16,32–36] However, their use as a general index for the *quantification* of the diradical character has not yet been fully explored. This is the main goal of the present work.

Computational Details

Since nondynamical correlation is essential to describe the low-spin components of diradicals, the use of a multireference method is mandatory. All wave functions for the molecular systems studied have been obtained at the CASSCF level with the cc-pVTZ basis set, unless otherwise indicated. For the simple diradical model systems the STO-3G basis set in combination with CASSCF or UHF levels of theory has been used instead. The first- and second-order density matrices have been obtained using a modified version of Gaussian03[37] and an auxiliary program[38] that reads and processes the CASSCF outputs. All local spin components are given in atomic units. All calculations have been carried out at the geometrical structure of the molecules optimized at the current level of theory, unless otherwise indicated. The local spin analysis has been performed with the program APOST-3D.[39] For this work we have make use of the atomic domains provided by the recently introduced topological fuzzy Voronoi cells (TFVC) scheme.[40] It is a fuzzy-atom based alternative[41] to Bader's QTAIM domains that produces very similar results with much less computational effort.[40]

Results and discussion

Local spin vs density of effectively unpaired electrons

In Figures 1 and 2 we plot the evolution of the indices of Eqn. 1-4 for diradical character for a simple model system, namely the dissociation of a singlet H_2 into two doublet H atoms described with minimal basis at the UHF and FCI levels of theory, respectively. The values of the number of effectively unpaired electrons averaged over one of the H atom, N_D^H and the local spin, $\langle \hat{S}^2 \rangle_H$, are also included. The later has been rescaled to vary from 0 to 1 for better comparison.

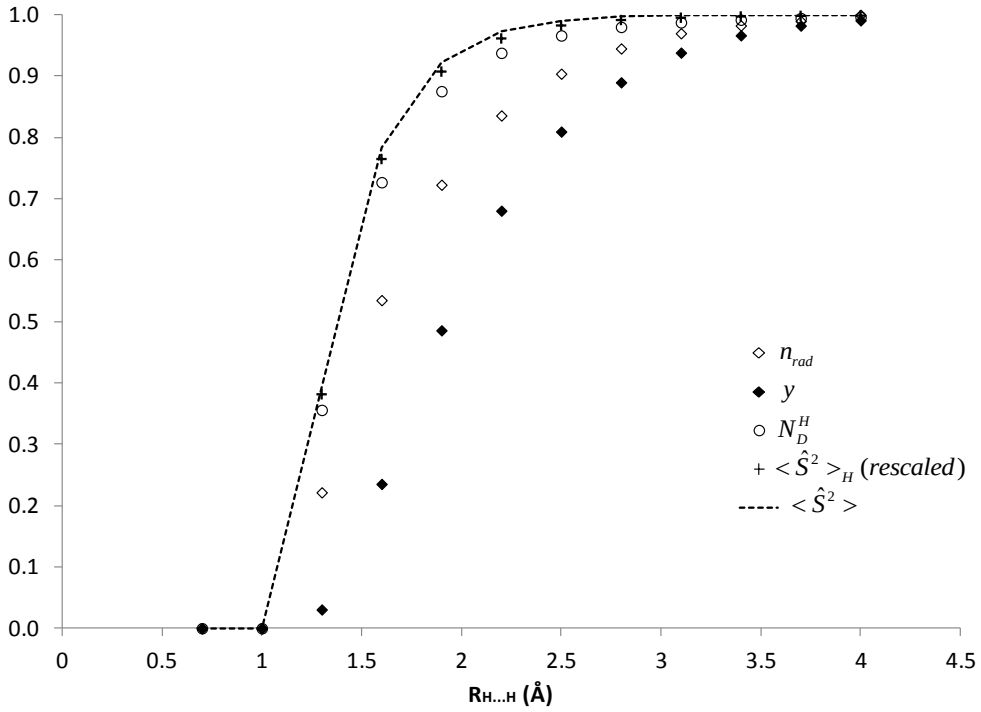

Figure 1: Indices for diradical character along dissociation of H_2 model system at the UHF level of theory. $\langle \hat{S}^2 \rangle_H$ values have been rescaled (see text).

For a single-determinant wave function the diradical character is exactly zero for

all indices when no BS solution exists (see Figure 1). As the H-H distance stretches and a BS solution is found, the diradical character monotonically increases in all cases. At large distances all indices tend to 1, indicating a perfect diradical. For intermediate distances, the index y from Eqn.(2) seems to underestimate the extent of diradical character with respect to the other indicators. Both the local spin and the number of effectively unpaired electrons closely follow the value of $\langle \hat{S}^2 \rangle$. For this model system the n_{rad} index is equivalent to the occupation of the LUMO orbital (n_{LUMO}).

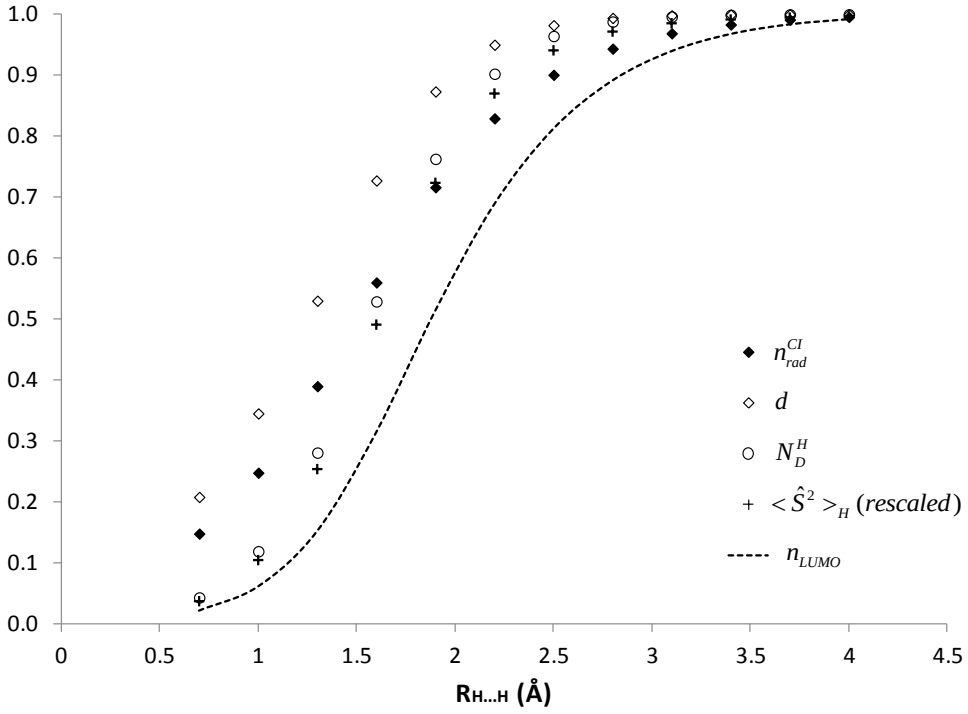

Figure 2: Indices for diradical character along dissociation of H_2 model system at the FCI level of theory. $\langle \hat{S}^2 \rangle_H$ values have been rescaled

For correlated wave functions the diradical character predicted by the different indices is always different from zero. The d index of Eqn. (4) yields a 20% of

diradical character for H_2 at the equilibrium distance (0.74 Å), which is probably somewhat too large. The local spin and number of effectively unpaired electrons on the H atoms give a similar and much smaller diradical character, and the index given in eqn. (3) lies in between. In this case, the occupation of the LUMO (antibonding) orbital consistently yields a smaller diradical character than the other indices. At intermediate atomic distances the differences between all indices are smaller than in the case of the UHF-BS description.

It is worth to note that Clark and Davidson[32] also applied their local spin formalism to the dissociation profile of diatomics such as H_2 at RHF, UHF and FCI levels of theory. In their method, the $\langle \hat{S}^2 \rangle_A$ value also tends to 3/4 at the dissociation limit, but for interatomic distances near equilibrium it tends to 3/8 of the bond order. The non-vanishing local spin contributions for a genuinely diamagnetic system like H_2 at equilibrium distance difficult their use as indicators of diradical character.

We have just seen that both the number of effectively unpaired electrons and the local spin analysis quantify in a similar manner the diradical character for a simple model system. Indeed, several studies have shown that both are very useful tools for the characterization of the spin distribution in actual molecular systems, in particular for singlet states.[15,16,32–35] However, $u(\vec{r})$ also exhibits some unattractive peculiarities. First of all, the upper bound for N_D was found to be $2N$, where N is the total number of electrons. Thus, the number of effectively unpaired electrons may be larger than the actual number of electrons.[42] This unphysical upper bound hinders the use of N_D as an *absolute* index for radical character. Another rather puzzling result was found in the dissociation of O_2 in its $^3\Sigma_g^-$ ground state into two triplet ^3P oxygen atoms. Staroverov and Davidson[42] obtained a value of N_D=5 at the dissociation limit, *i.e.,* each O atom carries an average of 2.5 unpaired electrons, instead of the expected value of 2 for an isolated triplet. It is worth mentioning that this finding motivated an alternative definition of $u(\vec{r})$ by Head-Gordon,[43] although not without controversy.[44,45]

We have further explored this paradigmatic system by considering for a number of different electronic states the dissociation of O_2 into two O atoms. In the dissociation limit one can have either two radical centers with two unpaired electrons each (when the O_2 dissociates into two triplet ^3P oxygen atoms), or no spin centers at all when it dissociates into two ^1D singlet O atoms. Note that neither situations can be

described with eqns. (1) to (4). In Table 1 we collect the values of N_D and local spin on the O atoms upon dissociation for several molecular (and atomic) electronic states. The wave functions have been obtained at the CASSCF(8,6)/6-31G* level of theory.

Table 1: Number of effectively unpaired electrons (N_D) and local spin values for the O atoms, $\langle \hat{S}^2 \rangle_{\mathrm{O}}$, at the dissociation limit of several O_2 molecular electronic states.

Molecular elec. state	N_D	$\langle \hat{S}^2 \rangle_{\mathrm{O}_1}$ / $\langle \hat{S}^2 \rangle_{\mathrm{O}_2}$	Atomic elec. state[a]
$^3\Sigma_g^-$	5	2 / 2	$^3P/^3P$
$^1\Delta_g$	4	2 / 2	$^3P/^3P$
$^1\Sigma_g^+$	5	2 / 2	$^3P/^3P$
$^1\Sigma_u^-$	5	2 / 2	$^3P/^3P$
$^3\Pi_u$	4	2 / 2	$^3P/^3P$
$^1\Pi_g$	5	2 / 2	$^3P/^3P$
$^1\Pi_u$	5	2 / 2	$^3P/^3P$
$^1\Delta_u$	5.33	0 / 0	$^1D/^1D$
$^1\Pi_u$	4.99	0 / 0	$^1D/^1D$

[a] Atomic electronic states at the dissociation limit

For the ground $^3\Sigma_g^-$ state, a value of N_D=5 is obtained upon dissociation into two triplet ^3P oxygen atoms, as already noted by Staroverov and Davidson.[42] However, this is not always the case. For instance, for the dissociation of the $^1\Delta_g$ and $^3\Pi_u$ states into two triplet O atoms, the expected N_D=4 value is recovered. The $^1\Delta_u$ state dissociates into two ^1D singlet oxygen atoms, but the N_D value is 16/3, consistent with the uniform distribution of 8 electrons into 6 degenerate p orbitals. Thus, by looking at the N_D values at the dissociation limit one can not distinguish two triplet from two singlet oxygen atoms (in this case the distinction is evident from the energy values). Moreover, different N_D values can be obtained for a system consisting of two dissociated triplet oxygen atoms, depending on the overall electronic state. It is worth to note that using Head-Gordon's[43] alternative formulation one would obtain N_D=4 in all cases (in fact, as long as the natural occupations are greater or equal than 1).

On the other hand, the local spin values *always* yield the expected values for the dissociating oxygen atoms. Matito and Mayer[24] already reported proper asymp-

totics of the atomic local spin contributions for the lowest-lying triplet and singlet states. We have considered here the dissociation of five more molecular singlet and triplet states that dissociate into two ^3P oxygen atoms and in all cases $\langle \hat{S}^2 \rangle_{\mathrm{O}} = 2$ (see Table 1). For the states that dissociate into two singlet ^1D oxygen atoms, namely $^1\Delta_u$ and $^1\Pi_u$, the local spin analysis yields $\langle \hat{S}^2 \rangle_{\mathrm{O}} = 0$. The diatomic spin components also differentiate when the two oxygen triplets are coupled as a singlet, like in the $^1\Delta_g$ state for which $\langle \hat{S}^2 \rangle_{\mathrm{O,O}} = -2$, or as a triplet, like in the $^3\Sigma_g^-$ state, for which $\langle \hat{S}^2 \rangle_{\mathrm{O,O}} = -1$ is obtained.

Thus, the local spin analysis appears to be more suitable tool than the number of effectively unpaired electrons when it comes to the formal breaking of more than one bond. This is in essence because the $\langle \hat{S}^2 \rangle_A$ terms include contributions from the cumulant of the second order-density matrix, whereas the number of effectively unpaired electrons is obtained only from the first-order density matrix. Accordingly, our goal, which is the quantification of diradical character, will be better accomplished by making use of the descriptors obtained from the local spin analysis.

Quantification of diradical character in molecules

The spin distribution of diradical species has already been analyzed in the light of the number of effectively unpaired electrons and different local spin indicators. Typically studied examples are benzyne isomers.[26,46–48] Clark and Davidson analyzed their electronic structure making use of the density of effectively unpaired electrons[49] and also their local spin formalism.[32,33] The evolution of local spins[32] and the number of unpaired electrons[34] along reactive processes involving benzyne were also discussed in detail.

For the present work we have studied a number of diradical and diradicaloid species at equilibrium geometries. The species considered are depicted in Figure 3. For all of them we have performed the local spin analysis, but the results will not be discussed in detail here (for that we refer to the supporting information). Instead, we will focus essentially on the actual quantification of the diradical character. For this purpose, only the atomic contributions of the local spin analysis will be taken into account.

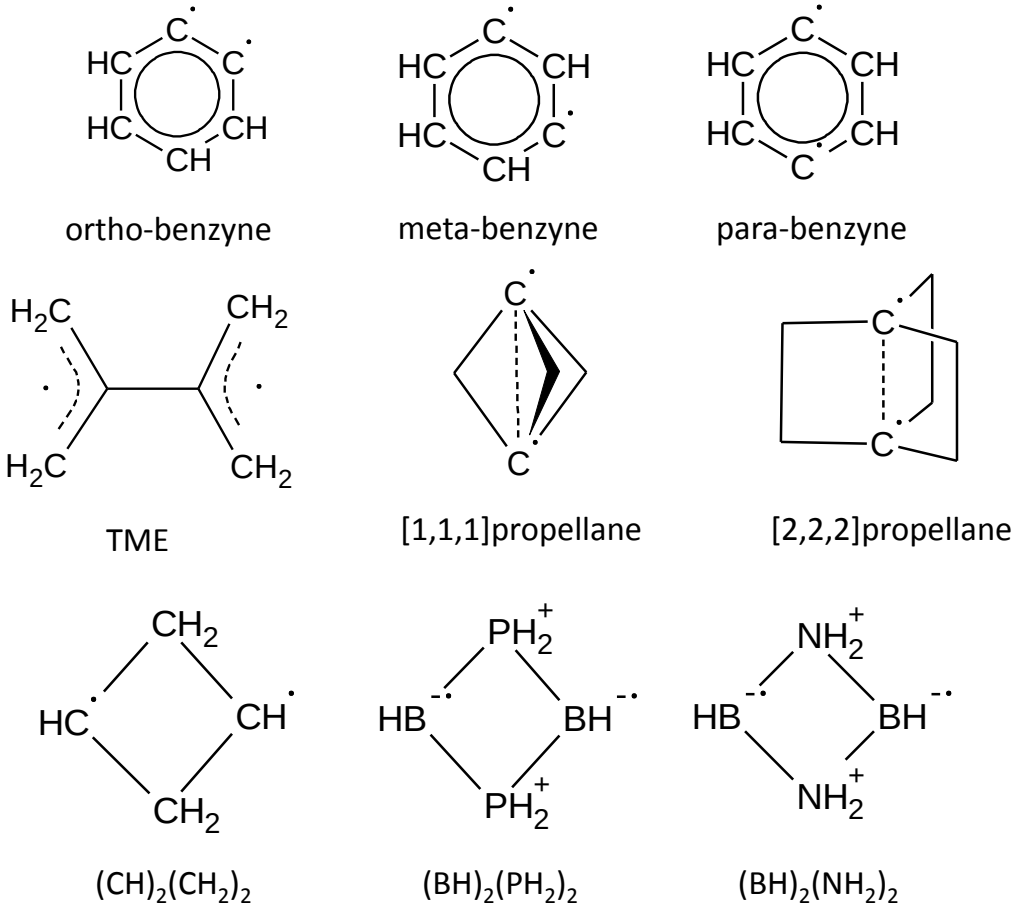

Figure 3: Diradicals and diradicaloids considered in this work

Most recently we have introduced as a general measure of k-radical character the following index

$$\Delta^{(k)} = \sqrt{\frac{\sum_A (\langle \hat{S}^2 \rangle_A - \langle \hat{S}^2 \rangle_A^{id})^2}{n}}, \tag{7}$$

where the $\langle \hat{S}^2 \rangle_A^{id}$ represent the atomic ideal values and n is the total number of

atoms/fragments considered in the local spin analysis (for most applications the hydrogen atoms can be safely ignored.)[31] Since the "ideal" value for the diatomic terms depends upon the particular electronic state (singlet, triplet, ...) and also the type of wave function (broken-symmetry vs multireference), the index only uses the atomic contributions of the local spin analysis. The main advantage of this index is that, contrary to most approaches in the literature, it is calculated in the same manner from both multireference and unrestricted single-determinant wave functions, and for any electronic state. The smaller the $\Delta^{(k)}$ value the closer the system is to a reference picture of k perfectly localized spin centers. So far, the use of eqn. 7 has been restricted to $k = 3$, for the quantification of the triradical character.[31]

For diradical or diradicaloid species in singlet-state one can calculate both $\Delta^{(0)}$ and $\Delta^{(2)}$ values. The former will measure average deviation from a nonradical closed-shell picture, and the latter will indicate the deviation from a perfect diradical. This provides a numerical criterion to identify diradicaloids as either diradicals or nonradical species.

Table 2: k-radical character index, $\Delta^{(k)}$, for a set of diradicals and diradicaloids. Values in parenthesis computed at the UB3LYP level of theory.

	Singlet		Triplet
Molecule	$\Delta^{(0)}$	$\Delta^{(2)}$	$\Delta^{(2)}$
ortho-benzyne	0.18	0.28	0.14
meta-benzyne	0.28	0.19	0.13
para-benzyne	0.49	0.12	0.12
TME	0.77	0.02	0.05
$(CH)_2(CH_2)_2$	0.41 (0.42)	0.12 (0.12)	0.12 (0.12)
$(BH)_2(PH_2)_2$	0.10 (0.00)	0.46 (0.53)	0.39 (0.34)
$(BH)_2(NH_2)_2$	0.17 (0.18)	0.41 (0.40)	0.38 (0.36)
$[1,1,1]$propellane	0.02	0.47	–
$[2,2,2]$propellane ($R_{C\text{-}C}$=1.536Å)	0.01	0.36	–
$[2,2,2]$propellane ($R_{C\text{-}C}$=1.988Å)	0.16	0.22	–
$[2,2,2]$propellane ($R_{C\text{-}C}$=2.532Å)	0.05	0.33	–

The computed $\Delta^{(0)}$ and $\Delta^{(2)}$ values for the species of Figure 3 are gathered on Table 2. The trends of the $\Delta^{(0)}$ and $\Delta^{(2)}$ values along the series of singlet ortho-,

meta- and para-benzyne are very illustrative. For ortho-benzyne $\Delta^{(0)} = 0.18$ and $\Delta^{(2)} = 0.28$, indicating a smaller deviation of this species with respect to a closed-shell picture. For meta-benzyne the situation is just the opposite, and the system is better identified as a diradical. For para-benzyne the $\Delta^{(2)}$ value is very small (0.12) and much smaller than the $\Delta^{(0)}$ one (0.49), which is consistent with a diradical picture. Thus, both $\Delta^{(0)}$ and $\Delta^{(2)}$ values are able to reproduce the assumed trend ortho $<$ meta $<$para of the diradical character of benzyne isomers.[32,46,47] For triplet states only the $\Delta^{(2)}$ values are meaningful. The values are quite small and do not differ too much from one isomer to another. The trend along the series is the same as for the singlet states, *i.e.,* , triplet para-benzyne is more diradical than ortho-benzyne.

Sometimes the spin delocalization hinders the recognition of the formal spin centers, like in the well-studied tetramethylenethane (TME) diradical.[50,51] At the CASSCF level of theory the lowest energy structure has D_2 symmetry, with a dihedral angle relating the two allyl moieties of 70.6 deg. The active space included 6 electrons and 6 orbitals (in the D_{2h} symmetry they correspond to the set of 6 π orbitals.) The results of the local spin analysis are discussed in detail in the supporting information. Essentially, the analysis reveals that TME is made up from two independent allyl radicals bonded by the central carbon atoms. The spin distribution among the atoms of the allyl fragments is very similar for both the singlet and triplet states.

Thus, for this molecule it is more appropriate to consider two allyl fragments in eq. 7, instead of all atoms separately. The local spin contribution of a molecular fragment is simply obtained by summing up all atomic and diatomic contributions of the atoms that form the molecular fragment. The $\Delta^{(2)}$ value taking the two allyl moieties as spin centers is very small (0.02) in the singlet state, and somewhat larger in the triplet (0.05) due to the enhanced delocalization of the spins between the two moieties. The $\Delta^{(0)}$ value is clearly too large to consider this system as a closed-shell species at all.

The distinction between a singlet diradical and a nonradical closed-shell species is sometimes not so evident. Diphosphadiboretanes and their analogues are some of the most controversial systems discussed in the literature. Scheschkewitz *et al.* [52]

reported several years ago a 1,3-diphospha-2,4-diboretane derivative singlet diradical that exhibited indefinite stability at room temperature. Several theoretical studies[3,8,17,53] followed that work, aimed at the quantification of the diradical character of this species and its analogues. We depict on Figure 3 some of these four-member ring diradicaloids. The species $(CH)_2(CH_2)_2$ corresponds to a planar transition state structure on the singlet potential energy surface of bicyclobutane (a triplet state lies ca. 2-3 kcal/mol lower in energy.) $(BH)_2(PH_2)_2$ is simplest diphosphadiboretane and $(BH)_2(NH_2)_2$ is a diaza-analogue of the former. This system is interesting because even though it exhibits a much shorter B-B distance (2.04 Å) than in diphosphadiboretane (2.60 Å), its diradical character was estimated to be smaller.[53]

We have studied these systems with an unrestricted single-determinant wave function (UB3LYP) for both their singlet and triplet states. We have also considered a single-point CASSCF(2,2) wave function at the UB3LYP optimized structures for comparison. For the singlet states, the atomic local spin values are very similar for CASSCF(2,2) and UB3LYP methods, provided a broken-symmetry solution is found for the latter (the local spin contributions are exactly zero for a restricted single-determinant wave function, as in $(BH)_2(PH_2)_2$). Remarkably, a broken-symmetry wave function that yields a wrong value of $\langle \hat{S}^2 \rangle$ does seem to provide appropriate atomic $\langle \hat{S}^2 \rangle_A$ contributions. The flaw of the broken-symmetry solution is found on the diatomic spin-spin interactions between the local spin centers: the UB3LYP values are significantly smaller than the CASSCF(2,2) ones. Yet, the negative sign still indicates the antiparallel arrangement of the local spins (see supporting information). For triplet states the local spin analysis yields very similar one- and two-center contributions for both methods. This is not surprising since with a CASSCF(2,2) approach the $m_S = |S|$ state is described by a ROHF wave function. Therefore, since the indices of eqn. 7 use only the atomic local spin contributions, their values for a broken-symmetry and a CASSCF wave function will be very similar.

We find that singlet $(CH)_2(CH_2)_2$ is best described as a diradical. The $\Delta^{(2)}$ value is similar to that of para-benzyne (0.12), whereas $\Delta^{(0)}$ is much larger. These values are indeed almost the same for CASSCF(2,2) and UB3LYP wave functions. The $\Delta^{(0)}$ value for $(BH)_2(PH_2)_2$ is trivially zero at the UB3LYP level of theory, as it corresponds to a restricted closed-shell solution. For CASSCF(2,2) the value

slightly increases to 0.10 but still is significantly smaller than the $\Delta^{(2)}$ value. Clearly, this species can not be considered a diradical, in agreement with Jung *et al.* .[53] For the diaza analogue, $\Delta^{(0)}$ increases to ca. 0.17 and $\Delta^{(2)}$ decreases to 0.41. Thus, $(BH)_2(NH_2)_2$ is more diradical than $(BH)_2(PH_2)_2$, but still it is best described as a closed-shell species. For triplet states, the local spin analysis reveals in the case of $(BH)_2(PH_2)_2$ and $(BH)_2(NH_2)_2$ that the four atoms of the ring exhibit similar but small local spin contributions (see supporting information). The large $\Delta^{(2)}$ values for the triplet states of $(BH)_2(PH_2)_2$ and $(BH)_2(NH_2)_2$ are thus consistent with the observed delocalized-spin picture.

Finally, the nature of the central C-C bond in strained systems such as propellanes has been subjected to debate in the literature for years. The formal picture of these species in the absence of this bond would be a diradical. However, the diradical character in the ground state has been ruled out in the case of [1,1,1]propellane by Wu et al.,[54] on the basis of a large vertical singlet-triplet gap (over 100 kcal/mol). The authors used a detailed Valence Bond analysis to classify the central C-C interaction as charge-shift bond. Lobayan et. al.[55] also analyzed the density of unpaired electrons and its topology for this species at the CISD level of theory and ruled out the presence of a 3c-2e bond. Yet, the overall number of unpaired electrons (N_D) they obtained at the CISD/6-31G* level of theory was quite significant (ca. 1.22).

We have performed the local spin analysis for [1,1,1]propellane and [2,2,2]propellane species at the CASSCF(10,10)/cc-pVTZ//UB3LYP/cc-pVTZ level of theory. The results are gathered on Table 3. For [1,1,1]propellane, the central C-C distance is 1.568 , very similar to that of ethane for the same level of theory (1.528). We have found that the local spin on the central C atoms is completely negligible. In fact it is even smaller than the local spin con the C atoms of ethane described at the same level of theory (0.009 and 0.018, respectively). Accordingly, the $\Delta^{(0)}$ value is very close to zero (0.018), as expected for a nonradical species. For this level of theory we obtain an overall N_D=0.46, a value significantly smaller than that obtained by Lobayan and in more agreement with a nonradical picture.

Table 3: Local spin on central C atoms, central C-C bond orders (see text) and total number of unpaired electrons for ethane and several propellane species at the CASSCF(10,10)/cc-pVTZ//UB3LYP/cc-pVTZ level of theory

Molecule	$R_{C\text{-}C}$ (Å)	$BO_{C\text{-}C}$ (fluct.)	$BO_{C\text{-}C}$ (exch.)	$\langle \hat{S}^2 \rangle_A$	N_D
ethane	1.528	0.78	0.91	0.018	0.318
[1,1,1]propellane	1.568	0.44	0.66	0.009	0.463
[2,2,2]propellane	1.536	0.77	0.87	0.028	0.095
	1.988[a]	0.33	0.43	0.313	1.082
	2.532	0.00	0.15	0.091	0.700

[a] Transition state structure.

The potential energy surface of [2,2,2]propellane was studied in detail by Davidson[56] with different levels of theory with the 6-31G* basis set. There are two similar minimum energy structures for the singlet state. In the most strained one, the central C-C distance is ca. 1.54 Å. Another minimum energy structure is found at a much longer C-C distance (2.54 Å). Both are connected by a transition state structure at an intermediate distance of ca 2 Å. Both UB3LYP and CASSCF(n,n) methods with n=2,4,8 yield similar structures and energetics. The strained minimum structure is about 5-10 kcal/mol higher in energy than the stretched one, and the barrier for the interconversion (from the strained structure) is about 15-20 kcal/mol.[56] The multireference average quadratic coupled cluster (MRAQCC) results obtained by Antol et. al.[57] with the same basis set were very similar to those reported by Davidson. It is worth to note that Davidson found that a low-energy broken-symmetry solution occurs from a C-C distance of ca. 1.7 Å at the UHF/6-31G* level, whereas for UB3LYP the broken-symmetry solution only exists between C-C distances of 1.9 to 2.3 Å.

We have optimized the three structures at the UB3LYP/cc-pVTZ level of theory. Only the transition state structure lead to a broken-symmetry solution. Then we carried out single-point energy calculations at the CASSCF(10,10)/cc-pVTZ level to perform the local spin analysis. The a_1', a_2" and two sets of e' and e" orbitals were included in the active space. For the strained minimum ($R_{C\text{-}C}$=1.536Å) the local spin in the central C atoms is again negligible (0.028), and so is the number

of unpaired electrons (N_D=0.095). The corresponding $\Delta^{(0)}$ value is similar to that obtained for [1,1,1]propellane.

In the stretched global minimum structure ($R_{C\text{-}C}$=2.532Å) there is no central C-C bond. The bond orders are 0.15 and 0.00 for the exchange and fluctuation formulations, respectively).[58] Also, the number of unpaired electrons is significantly larger than for the strained structure (N_D=0.70). Yet, the local spin on the central C atoms is still very small (0.091). The $\Delta^{(2)}$ value of 0.33 is too large to consider this species as a diradical at all, specially when compared with the value for $\Delta^{(0)}$ (0.05). Since the UB3LYP description of this species is spin-restricted, $\Delta^{(0)}$=0 by definition at this level of theory.

The transition state structure ($R_{C\text{-}C}$=1.988Å) does exhibit significant local spin in the central C atoms (0.31), as well as larger number of effectively unpaired electrons (N_D=1.08). The $\Delta^{(0)}$ and $\Delta^{(2)}$ values are 0.16 and 0.22, respectively. Thus, the diradicaloid character at the transition state is larger than that of the minimum energy structures, but the species is still best pictured as a nonradical.

Conclusions

The general quantification of diradical character from wave function analysis is shown to be a non-trivial task, particularly for singlet states. In this work we illustrate how the descriptors obtained from a local spin analysis can be used to define a general measure of the diradical character. Indices $\Delta^{(0)}$ and $\Delta^{(2)}$ quantify deviation from a nonradical and a perfect diradical picture, respectively. The method reproduces the expected trend ortho-benzyne<meta-benzyne<para-benzyne of diradical character, for both the singlet and the triplet states. Also, it is found that diphospadiboretane and its diaza-analogue are best described as closed-shell and delocalized-spin species in their singlet and triplet states, respectively. The analysis performed on strained propellanes also confirm their nonradical nature, even in the absence of the central C-C bond.

Acknowledgments

Financial help has been furnished by the Spanish MICINN Projects No. CTQ2011-23441/BQU. Financial support from MICINN and the FEDER fund (European Fund for Regional Development) was also provided by grant UNGI08-4E-003. Financial support from the Generalitat de Catalunya (SGR528 and Xarxa de Referència en Química Teòrica i Computacional) is also acknowledged. E.R-C. acknowledges support from the Spanish FPU program (Grant No. AP2008-01231) and from the EU under a Marie Curie Career Integration grant (PCI09-GA-2011-294240).

References

[1] L. Salem, *Angew. Chem. Int. Ed. Engl.*, 1972, **11**, 92.

[2] A. Rajca, *Chem. Rev.*, 1994, **94**, 871–893.

[3] Y. Jung and M. Head-Gordon, *Chem. Phys. Chem.*, 2003, **4**, 522–525.

[4] L. Noodleman, *J. Chem. Phys.*, 1981, **74**, 5737.

[5] L. Noodleman and E. R. Davidson, *J. Chem. Phys.*, 1986, **109**, 131.

[6] L. Salem and C. Rowland, *Angew. Chem.*, 1972, **84**, 86.

[7] V. Bachler, G. Olbrich, F. Neese and K. Wieghardt, *Inorg. Chem.*, 2002, **41**, 4179–4193.

[8] M. Seierstad, C. R. Kinsinger and C. J. Cramer, *Angew. Chem. Int. Ed. Engl.*, 2002, **20**, 41.

[9] X. Lopez, F. Ruipérez, M. Piris, J. M. Matxain and J. M. Ugalde, *Comp. Phys. Com.*, 2011, **12**, 1061–1065.

[10] P. Rivero, C. A. Jiménez-Hoyos and G. E. Scuseria, *J. Phys. Chem. B*, 2013, **117**, 12750–12758.

[11] K. Kamada, K. Ohta, A. Shimizu, T. Kubo, R. Kishi, H. Takahashi, E. Botek, B. Champagne and M. Nakano, *J. Phys. Chem. Lett.*, 2010, **1**, 937–940.

[12] D. Herebian, K. E. Wieghardt and F. Neese, *J. Am. Chem. Soc.*, 2003, **125**, 10997–11005.

[13] F. Plasser, H. Pašalić, M. H. Gerzabek, F. Libisch, R. Reiter, J. Burgdörfer, T. Müller, R. Shepard and H. Lischka, *Angew. Chem. Int. Ed. Engl.*, 2013, **52**, 2581–2584.

[14] K. Takatsuka, T. Fueno and K. Yamaguchi, *Theor. Chim. Acta (Berlin)*, 1978, **48**, 175–183.

[15] V. N. Staroverov and E. R. Davidson, *J. Am. Chem. Soc.*, 2000, **122**, 186–187.

[16] V. N. Staroverov and E. R. Davidson, *J. Am. Chem. Soc.*, 2000, **122**, 7377–7385.

[17] M.-J. Cheng and C.-H. Hu, *Mol. Phys.*, 2003, **101**, 1319–1323.

[18] L. Lain, A. Torre, R. C. Bochicchio and R. Ponec, *Chem. Phys. Lett.*, 2001, **346**, 283–287.

[19] I. Mayer, *Int. J. Quant. Chem.*, 1986, **29**, 73–84.

[20] I. Mayer, *Int. J. Quant. Chem.*, 1986, **29**, 477–483.

[21] A. E. Clark and E. R. Davidson, *J. Chem. Phys.*, 2001, **115**, 7382–7392.

[22] M. Reiher, *Faraday Discuss.*, 2006, **135**, 97–124.

[23] I. Mayer, *Chem. Phys. Lett.*, 2007, **440**, 357–359.

[24] I. Mayer and E. Matito, *Phys. Chem. Chem. Phys.*, 2010, **12**, 11308–11314.

[25] D. R. Alcoba, A. Torre, L. Lain and R. C. Bochicchio, *J. Chem. Theory Comput.*, 2011, **7**, 3560–3566.

[26] E. Ramos-Cordoba, E. Matito, P. Salvador and I. Mayer, *Phys. Chem. Chem. Phys.*, 2012, **14**, 15291–15298.

[27] E. Ramos-Cordoba, E. Matito, I. Mayer and P. Salvador, *J. Chem. Theory Comput.*, 2012, **8**, 1270–1279.

[28] C. Herrmann, L. Yu and M. Reiher, *J. Comput. Chem.*, 2006, **27**, 1223–1239.

[29] C. Herrmann, M. Reiher and B. A. Hess, *J. Chem. Phys.*, 2005, **122**, 034102.

[30] E. Ramos-Cordoba, P. Salvador and M. Reiher, *Chem. Eur. J.*, 2013, **19**, 15267–15275.

[31] E. Ramos-Cordoba and P. Salvador, *J. Chem. Theory Comput.*, 2014, DOI: 10.1021/ct401009p.

[32] A. E. Clark and E. R. Davidson, *J. Phys. Chem. A*, 2002, **106**, 6890–6896.

[33] A. E. Clark and E. R. Davidson, *Mol. Phys.*, 2002, **100**, 373–383.

[34] A. Clark, E. Davidson and J. Zaleski, *J. Am. Chem. Soc.*, 2001, **123**, 2650–2657.

[35] A. Clark and E. Davidson, *Phys. Chem. Chem. Phys.*, 2007, **9**, 1881–1894.

[36] J. Oliva, D. Alcoba, L. Lain and A. Torre, *Theor. Chem. Acc.*, 2013, **132**, 1–6.

[37] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03, Revision C.02*, Gaussian, Inc., Pittsburgh, PA, 2003.

[38] E. Matito and F. Feixas, *DMN program*, 2009, University of Girona (Spain) and University of Szczecin (Poland).

[39] P. Salvador and E. Ramos-Cordoba, *APOST-3D program*, 2012, Universitat de Girona (Spain).

[40] P. Salvador and E. Ramos-Cordoba, *J. Chem. Phys.*, 2013, **139**, 071103.

[41] E. Matito, M. Solà, P. Salvador and M. Duran, *Faraday Discuss.*, 2007, **135**, 325–345.

[42] V. N. Staroverov and E. R. Davidson, *Chem. Phys. Lett.*, 2000, **330**, 161–168.

[43] M. Head-Gordon, *Chem. Phys. Lett.*, 2003, **372**, 508–511.

[44] R. C. Bochicchio, A. Torre and L. Lain, *Chem. Phys. Lett.*, 2003, **380**, 486–487.

[45] M. Head-Gordon, *Chem. Phys. Lett.*, 2003, 488–489.

[46] P. G. Wenthold, R. R. Squires and W. Lineberger, *J. Am. Chem. Soc.*, 1998, **120**, 5279–5290.

[47] F. De Proft, P. von Ragué Schleyer, J. H. van Lenthe, F. Stahl and P. Geerlings, *Chem. Eur. J.*, 2002, **8**, 3402–3410.

[48] E. B. Wang, C. A. Parish and H. Lischka, *J. Chem. Phys.*, 2008, **129**, 044306.

[49] A. Clark and E. Davidson, *J. Am. Chem. Soc.*, 2001, **123**, 10691–10698.

[50] P. Nachtigall and K. D. Jordan, *J. Am. Chem. Soc.*, 1992, **114**, 4743–4747.

[51] M. Filatov and S. Shaik, *J. Phys. Chem. A*, 1999, **103**, 8885–8889.

[52] D. Scheschkewitz, H. Amii, H. Gornitzka, W. Schoeller, D. Bourissou and G. Bertrand, *Science*, 2002, **295**, 1880–1881.

[53] Y. Jung and M. Head-Gordon, *J. Phys. Chem. A*, 2003, **107**, 7475–7481.

[54] W. Wu, J. Gu, J. Song, S. Shaik and P. C. Hiberty, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 1407–1410.

[55] R. Lobayan and R. Bochicchio, *Chem. Phys. Lett.*, 2013, **557**, 154–158.

[56] E. Davidson, *Chem. Phys. Lett.*, 1998, **284**, 301–307.

[57] I. Antol, M. Eckert-Maksić, H. Lischka and Z. B. Maksić, *Eur. J. Org. Chem.*, 2007, **2007**, 3173–3178.

[58] J. G. Ángyán, E. Rosta and P. R. Surján, *Chem. Phys. Lett.*, 1999, **299**, 1–8.

SUPPLEMENTARY INFORMATION

Diradical character from the local spin analysis

Eloy Ramos-Cordoba, Pedro Salvador
Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química,
University of Girona, 17071 Girona, Spain
e-mail: Pedro.Salvador@udg.edu

1 The local spin analysis

In the local spin analysis, the $\langle \hat{S}^2 \rangle_A$ values indicate and quantify the presence of local spin *within* the molecule, namely on atom/fragment A. The magnitude and sign of the diatomic contributions $\langle \hat{S}^2 \rangle_{AB}$ with $B \neq A$ inform about the nature of the couplings between these local spins. The physical interpretation of the $\langle \hat{S}^2 \rangle_A$ and particularly $\langle \hat{S}^2 \rangle_{AB}$ values is somewhat intricate, and has been recently discussed in detail in several papers.[1,2] We provide here a brief account of its general characteristics for ideal systems.

When two perfectly localized spins are coupled as a singlet, a proper multireference wave function is needed to account for the spin properties of the system. In that case, the local spin analysis yields $\langle \hat{S}^2 \rangle_A = 3/4$ and $\langle \hat{S}^2 \rangle_{AB} = -3/4$, which account for the expected overall $\langle \hat{S}^2 \rangle = 0$. The $\langle \hat{S}^2 \rangle_A$ value is consistent with the corresponding $\langle \hat{S}^2 \rangle = s(s+1)$ value for the isolated one-electron system and the negative sign of $\langle \hat{S}^2 \rangle_{AB}$ indicates that the two local spins are coupled as a singlet (entangled). With a single-determinant broken symmetry description of the same system one would obtain similarly $\langle \hat{S}^2 \rangle_A = 3/4$, but now $\langle \hat{S}^2 \rangle_{AB} = -1/4$, for an overall value of $\langle \hat{S}^2 \rangle = 1$. Both $\langle \hat{S}^2 \rangle_A$ and $\langle \hat{S}^2 \rangle_{AB}$ monotonically decrease as the two spins become more amd more delocalized. In the limiting case of a closed-shell single-determinant description, all local spin contributions exactly vanish.

If the two perfectly localized spins on centers A and B are parallel, the local spin analysis would yield $\langle \hat{S}^2 \rangle_A = 3/4$ and $\langle \hat{S}^2 \rangle_{AB} = 1/4$, for an overall value of $\langle \hat{S}^2 \rangle = 2$, as expected for a triplet. The one-center term is again consistent with a one-electron system, and the positive sign of the diatomic contribution now indicates that the local spins are parallel. In a parallel-spins situation, if the two spins are not perfectly localized the value of $\langle \hat{S}^2 \rangle_{AB}$ decreases, but that of $\langle \hat{S}^2 \rangle_A$ increases. The latter is an indication of partial triplet character on the given center.

2 Local spin analysis of diradicals and diradicaloids

Tetramethyleneethane (TME) is a well-studied diradical.[3,4] This molecule has been optimized under three different symmetry constrains, namely D_2, D_{2h} and D_{2d} at the CASSCF level of theory. The active space included 6 electrons and 6 orbitals (in the D_{2h} symmetry they correspond to the set of 6 π orbitals.) The lowest energy structure has D_2 symmetry, with a C_1-C_2-C_4-C_5 dihedral angle (α) of 70.6 deg (see Figure 1).

The results of the local spin analysis are gathered on Table 1. The atomic and diatomic spin components are almost independent of the rotation with respect to the central C-C bond. The vertical singlet-triplet gaps, sometimes used to assess the radical character,[4] are not too different for the D_2, D_{2h}, and D_{2d} structures (-1.33, -3.91, and -2.05 kcal/mol, respectively). We will focus on the results obtained for the global minimum (D_2 symmetry). The main spin centers are C_1 and the symmetry equivalent C_3, C_5, and C_6, with $\langle \hat{S}^2 \rangle_C$ values of 0.35 and 0.36 for the singlet and triplet states (the local spin involving the H atoms is negligible). The diatomic spin terms $\langle \hat{S}^2 \rangle_{C_1,C_2}$ and $\langle \hat{S}^2 \rangle_{C_1,C_3}$ (and their symmetry equivalents) also equal in both electronic states. The sign of these spin contributions indicates the alternation of the spins within each allyl fragment, as indicated in Figure 1. The main differences between the local spin distribution of the singlet and triplet states are found in the diatomic terms involving the C atoms on the different allyl fragments. In the singlet state, the atoms 1, 3, and 4 have parallel spins, as indicated by the sign of $\langle \hat{S}^2 \rangle_{C_1,C_3}$ and $\langle \hat{S}^2 \rangle_{C_1,C_4}$ terms, whereas centers 2, 5, and 6 exhibit antiparallel arrangement with respect to them.

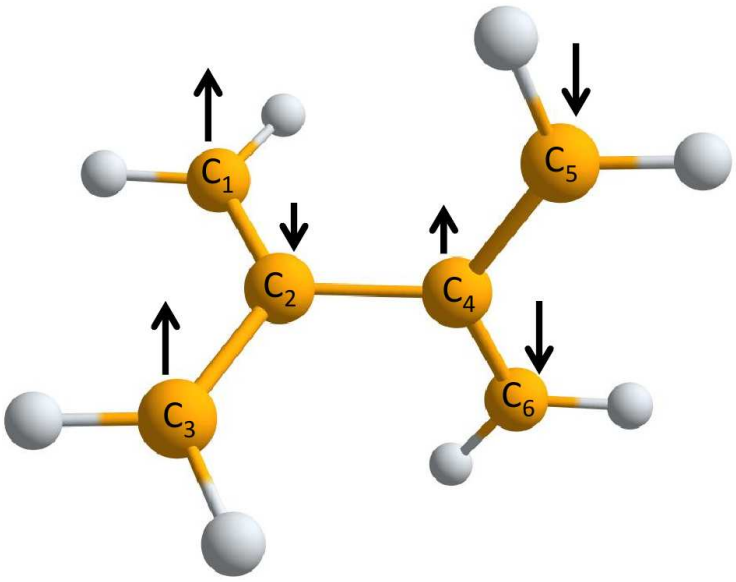

Figure 1: Numbering scheme of tetrathyleneethane (TME) and local spin distribution of the singlet state.

The diatomic terms $\langle \hat{S}^2 \rangle_{C_1,C_5}$ present a rather small value of -0.17. However, this value accounts for most of the expected diatomic spin contribution of -3/4 between the two allyl centers, as there are four such diatomic contributions equivalent by symmetry. In the triplet state, the local spins on the C atoms of one of the allyl moieties are flipped with respect to the singlet state. The four symmetry-equivalent diatomic terms $\langle \hat{S}^2 \rangle_{C_1,C_5} = 0.05$ account for the spin-spin interactions between the allyl moieties.

The diatomic spin contribution between atoms C_2 and C_4 is almost zero. Moreover, the one- and two-center contributions involving atoms C_1, C_2 and C_3 are very similar to those observed for a single allyl radical.[5] In fact, summing up all one- and two-center contributions for all atoms of each allyl moiety gives a local spin on each fragment very close to 3/4 in all cases (see bottom of Table 1). Thus, the TME molecule can be regarded as a diradical made up from two independent allyl radicals bonded by the central carbon atoms. The local spin analysis unravels the spin distribution among all centers or fragments on the same footing both different electronic states, *i.e., even if no spin density exists*.

Table 1: Local spin analysis of the TME molecule for the singlet (S) and triplet (T) states of different geometries.

Atom/Atom pair	D_2 ($\alpha = 70.6$)		D_{2h} ($\alpha = 0$)		D_{2d} ($\alpha = 90$)	
Fragment/ Fragment pair	S	T	S	T	S	T
C_1	0.36	0.37	0.36	0.37	0.35	0.37
C_2	0.14	0.14	0.14	0.14	0.14	0.14
C_1,C_2	-0.09	-0.09	-0.08	-0.09	-0.09	-0.09
C_1,C_3	0.08	0.08	0.08	0.08	0.08	0.08
C_1,C_4	0.03	-0.01	0.03	0.00	0.03	-0.01
C_1,C_5	-0.17	0.05	-0.17	0.04	-0.17	0.05
C_1,C_6	-0.18	0.05	-0.19	0.04	-0.17	0.05
C_2,C_4	-0.01	-0.01	-0.02	-0.02	-0.01	0.00
allyl	0.77	0.80	0.79	0.81	0.76	0.79
$allyl_1$,$allyl_2$	-0.77	0.20	-0.79	0.19	-0.76	0.21

Diphosphadiboretanes and their analogues are some of the most controversial diradicaloid systems discussed in the literature. For these systems we have carried out the local spin analysis with both an unrestricted single-determinant (UB3LYP) and a CASSCF(2,2) wave function for comparision, and for both their singlet and triplet states. The results are gathered on Table 2.

Table 2: Local spin analysis of four-member ring diradicaloids for different spin states and levels of theory

Molecule	Atom/Atom pair	$\langle\hat{S}^2\rangle_A/\langle\hat{S}^2\rangle_{AB}$			
		CASSCF		UB3LYP	
		Singlet	Triplet	Singlet	Triplet
$(CH)_2(C'H_2)_2$	C	0.58	0.59	0.59	0.59
	C'	0.05	0.05	0.05	0.06
	C-C	-0.46	0.15	-0.16	0.16
$(BH)_2(PH_2)_2$	B	0.10	0.27	0.00	0.32
	P	0.09	0.26	0.00	0.22
	B-B	-0.04	0.03	0.00	0.05
$(BH)_2(NH_2)_2$	B	0.19	0.28	0.20	0.31
	N	0.14	0.25	0.15	0.24
	B-B	-0.07	0.03	-0.03	0.05

The local spin values (atomic terms) for singlet states are very similar for CASSCF(2,2) and UB3LYP methods, provided a broken-symmetry solution is found for the latter. For singlet states, the species with a larger local spin contributions is $(CH)_2(CH_2)_2$. The value for the CASSCF(2,2) wave function (0.58) is not too far from that expected for a perfectly localized electron (3/4). In $(BH)_2(NH_2)_2$ the local spin is significantly smaller (0.19), which should indicate a much weaker diradical character. In the case of the diphosphadiboretane species the local spin is almost negligible (0.10), consistent with a residual diradical character.

For triplet states the local spin analysis yields very similar one- and two-center contributions for both methods. This is not surprising since with a CASSCF(2,2) approach the $m_S = |S|$ state is described by a ROHF wave function. The atomic contributions of $(CH)_2(CH_2)_2$ are essentially the same as in the singlet state. The diatomic term involving the two main local spin centers is now positive, indicating parallel arrangement of the spins. In the case of $(BH)_2(PH_2)_2$ and $(BH)_2(NH_2)_2$ the four atoms of the ring exhibit similar but small contributions. In fact, the sum of the terms reported on Table 2 is still far from the overall $\langle \hat{S}^2 \rangle \approx 2$ value. This is because the hydrogen atoms (omitted thus far) exhibit small but significant contributions of ca. 0.05-0.10. Thus, these species do not exhibit significant spin centers and the molecular spin is delocalized over all atoms.

References

[1] E. Ramos-Cordoba, P. Salvador and M. Reiher, *Chem. Eur. J.*, 2013, **19**, 15267–15275.

[2] E. Ramos-Cordoba and P. Salvador, *J. Chem. Theory Comput.*, 2014, DOI: 10.1021/ct401009p.

[3] P. Nachtigall and K. D. Jordan, *J. Am. Chem. Soc.*, 1992, **114**, 4743–4747.

[4] M. Filatov and S. Shaik, *J. Phys. Chem. A*, 1999, **103**, 8885–8889.

[5] E. Ramos-Cordoba, E. Matito, P. Salvador and I. Mayer, *Phys. Chem. Chem. Phys.*, 2012, **14**, 15291–15298.

5.3 Characterization and quantification of polyradical character

Ramos-Cordoba, E.; Salvador, P. “Characterization and Quantification of Polyradical Character”. *Journal of Chemical Theory and Computation*. 2014, 10, 634–641

http://dx.doi.org/10.1021/ct401009p

http://pubs.acs.org/doi/abs/10.1021/ct401009p

Received: November 22, 2013

Published: January 16, 2014

Abstract

The decomposition of [$\hat{S}2$] into atomic and diatomic contributions (local spin analysis) is used to detect and quantify the polyradical character of molecular systems. A model triradical system is studied in detail, and the local spin analysis is used to distinguish several patterns of local spin distributions and spin–spin interactions that can be found for different electronic states. How close a real molecular system is to an ideal system of k perfectly localized spin centers is utilized to define a measure of its k-radical character. The spin properties and triradical character of the lowest-lying electronic states of a number of all σ, all π, and σ–π organic triradicals are discussed in detail. The local spin contributions exhibit good correlation with experimental triradical stabilization energies.

Chapter 6

Effective atomic orbitals: developments and applications

6.1 The atomic orbitals of the topological atom

THE JOURNAL OF CHEMICAL PHYSICS **138**, 214107 (2013)

The atomic orbitals of the topological atom

Eloy Ramos-Cordoba,[1] Pedro Salvador,[1,a)] and István Mayer[2]
[1]*Institute of Computational Chemistry and Catalysis and Department of Chemistry, University of Girona, 17071 Girona, Spain*
[2]*Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, P.O. Box 17, H-1525 Budapest, Hungary*

(Received 22 March 2013; accepted 13 May 2013; published online 5 June 2013)

The effective atomic orbitals have been realized in the framework of Bader's atoms in molecules theory for a general wavefunction. This formalism can be used to retrieve from any type of calculation a proper set of orthonormalized numerical atomic orbitals, with occupation numbers that sum up to the respective Quantum Theory of Atoms in Molecules (QTAIM) atomic populations. Experience shows that only a limited number of effective atomic orbitals exhibit significant occupation numbers. These correspond to atomic hybrids that closely resemble the core and valence shells of the atom. The occupation numbers of the remaining effective orbitals are almost negligible, except for atoms with hypervalent character. In addition, the molecular orbitals of a calculation can be exactly expressed as a linear combination of this orthonormalized set of numerical atomic orbitals, and the Mulliken population analysis carried out on this basis set exactly reproduces the original QTAIM atomic populations of the atoms. Approximate expansion of the molecular orbitals over a much reduced set of orthogonal atomic basis functions can also be accomplished to a very good accuracy with a singular value decomposition procedure. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4807775]

INTRODUCTION

The concept of *atom in a molecule* has always craved for a proper definition. However, we are lacking a single, unambiguous one. Instead, over the last decades a number of schemes or formalisms have been devised to identify the atom *within* a molecule.[1–10] Probably, any of such schemes considers the nucleus as part of the atom, so the differences always arise in how the electron population distributed in the physical space (or in the Hilbert-space) is subdivided into atomic shares.

Within the LCAO approach, the atom may be identified with the subspace of the basis functions attached to it. Such approach leads to the so-called Hilbert-space analyses,[11] such as the classical Mulliken[1] or Löwdin[2] population analysis of the density. Despite their simplicity, Hilbert-space analyses have been criticized by their restricted applicability (the use of atom-centered basis functions is necessary) and their notable basis set dependence. The latter represents a true flaw when using extended basis sets including diffuse functions.[12] Alternative population analyses based upon occupation numbers[13,14] carried out onto an AO basis set different from the extended one minimize the basis set dependence.

A different strategy is to subdivide the physical three-dimensional (3D) space into atomic regions or domains, which represent (together with the nucleus) the atom. These domains may be defined disjoint, like in Bader's atoms in molecule theory[4] (often referred to as QTAIM—"quantum theory of atoms in molecules"), or may be allowed to overlap, like in the different flavors of "fuzzy" atoms.[3,5–10] The 3D space formalism represents a perfect counterpart of the Hilbert-space analysis. Indeed, by introducing a proper mapping,[15,16] one can find a one-to-one correspondence between the expressions of quantities such as atomic populations and bond orders,[5] energy components,[17–19] or local spins,[20] obtained in these two frameworks. Of course, the actual values of these quantities derived from one or another formalism differ. Moreover, the QTAIM analysis may yield domains with so-called non-nuclear attractors, which correspond to regions of the space with no nucleus associated. Often, the appearance of a non-nuclear attractor is an artifact of the basis set applied, such as in the case of acetylene.[21,22] In such cases, it may be worth to use another partitioning of the space.

When looking for an *atom in a molecule*, obviously we are not merely interested in a subdivision of the 3D space into atomic volumes, but rather in assigning different physical quantities to the individual atoms (or their groups). Any physical quantity is expressed as the expectation value of one-electron (two-electron) operators, and it can also be written in terms of the integral over the space of the appropriate one-electron (two-electron) density functions. Hence, the subdivision of the space into atomic regions naturally leads to the decomposition of different physical quantities into atomic (diatomic) terms that can be considered the *effective* atomic (or diatomic) values of that physical quantity within the molecule. This has been extensively accomplished within the QTAIM framework, partly due to the special properties provided by the zero-flux condition, such as the local fulfillment of the virial theorem.[4] However, it has been shown that one can obtain quantities such as atomic populations and valences,[5] energy components,[17–19] or local spins[20] for essentially any atom in molecule definition. Similarly, the atom in

a)E-mail: pedro.salvador@udg.edu

the molecule can also be characterized from the analysis of the density matrix.[23–25]

Probably, the most appropriate entities that serve to characterize the state of the atom within the molecule are the so-called effective atomic orbitals ("effective AOs").[26–31] In this approach, one obtains for each atom a set of orthogonal atomic hybrids and their respective occupation numbers usually adding up to the *net* population of the atom. A remarkable feature of this scheme is that, for a given molecular system, irrespective of the basis set size applied in the calculation, one in practice obtains the same number of significantly populated atomic hybrids. This number has always been found equal to the number of the orbitals in a classical minimal basis, which means that in this manner one obtains "effective minimal basis sets" by performing *a posteriori* analyses of the calculations (special provisions should be made in the case of hypervalent systems). This happens to be the case even if no atom-centered basis functions are used at all.[31] These atomic hybrids closely mimic the core and valence shells of the atom, as anticipated on the basis of classical notions of electron configuration of the atom/fragment within the molecule.

This formalism was first introduced nearly two decades ago in the framework of Hilbert-space analysis,[27] and then it was generalized for an arbitrary Hermitian bilinear "localization functional."[28] It was applied to the case of "fuzzy" atoms a few years ago.[30] It had also been formulated in the context of Bader's QTAIM,[29] but until now that version was never actually realized.

When the atoms are associated with non-overlapping domains, as is the case of Bader's QTAIM, the "effective AOs" have special properties that make them very appealing from both conceptual and practical points of view, as will be shown later on. The first feature is that the "effective AOs" associated to different atoms are also orthogonal, as a direct consequence of not sharing at all the physical space. Also, because the atoms are not allowed to overlap, the atom's net and gross populations are equal. In other words, the sum of the occupation numbers of the "effective AOs" of a given atom is equal to its atomic population.

It is fair to note that the formalism of the domain-averaged Fermi hole analysis[32,33] (DAFH) also produces orbital functions in the framework of QTAIM (or other AIM schemes). With the DAFH analysis, one first obtains a set of domain orbitals and orbitals occupancies, that have their origin in the average of the exchange-correlation density over a space domain, typically the union of several atomic domains. These orbitals are then localized via a non-unitary isopycnic transformation,[34] leading to a new set of objects that are interpreted in terms of bonding orbitals, lone pairs, and dangling valences. Even though their origin is different, the DAFH and "effective AO" analyses share some similarities, particularly when the DAFH analysis is carried out over a single atomic domain. In fact, in the restricted single-determinant case the orbitals before isopycnic localization, sometimes denoted as domain natural orbitals,[33] are exactly the same as the original "effective AOs."[29]

The special features of the "effective AOs" in the framework of QTAIM (or any disjoint partition of the space in general) suggests that they could be used as (numerical) atomic basin-centered orthogonal basis set, in which the actual molecular orbitals of the molecule can be expanded. Furthermore, since the "effective AOs" can be obtained even if no atomic basis functions are used at all[31] (e.g., plane wave calculations), this formalism can be used to actually retrieve from such type of calculations a proper set of orthogonal atomic basis functions.

Of course, the fact that the molecular orbitals can be expressed as a linear combination of an alternative set of basis functions, even if it is orthonormal and has a reduced number of functions, adds no special chemical relevance *per se*. But, as we show in the "Theoretical Methods" section, in this particular basis, the Hilbert-space and 3D space analyses turn out to be fully equivalent numerically. Thus, the Mulliken population analysis of the density carried out on the basis of "effective AOs" exactly reproduces the original QTAIM atomic populations of the atoms.

In other words, these "effective AOs" appear to be the genuine atomic orbitals of Bader's theory, perhaps one of the few ingredients missing in QTAIM's toolbox.

THEORETICAL METHODS

Let us consider a system with n orthonormalized doubly occupied orbitals $\varphi_i(\vec{r})$, $i = 1, 2\ldots, n$, and the division of the 3D space into N_{at} disjunct atomic domains Ω_A defined, e.g., by Bader's "topological" QTAIM method. Let us for each atom A ($A = 1, 2, \ldots, N_{at}$) form the $n \times n$ Hermitian matrix $\mathbf{Q}^A$ with the elements

$$Q_{ij}^A = \int_{\Omega_A} \varphi_i^*(\vec{r})\varphi_j(\vec{r})dv. \tag{1}$$

Matrix $\mathbf{Q}^A$ is essentially the "atomic overlap matrix" in the basis of the MOs φ_i.

Furthermore, for each atom A we define the "intraatomic" part φ_i^A of every MO φ_i as

$$\varphi_i^A(\vec{r}) = \begin{cases} \varphi_i(\vec{r}) \text{ if } \vec{r} \in \Omega_A; \\ 0 \quad \text{if } \vec{r} \notin \Omega_A. \end{cases} \tag{2}$$

Thus,

$$Q_{ij}^A = \langle \varphi_i^A | \varphi_j^A \rangle, \tag{3}$$

i.e., $\mathbf{Q}^A$ is the overlap matrix of the orbitals φ_i^A.

Owing to the disjunct character of the atomic domains, one obviously has

$$\varphi_i(\vec{r}) = \sum_A^{N_{at}} \varphi_i^A(\vec{r}). \tag{4}$$

We diagonalize the Hermitian matrix $\mathbf{Q}^A$ by the unitary matrix $\mathbf{U}^A$

$$\mathbf{U}^{A\dagger}\mathbf{Q}^A\mathbf{U}^A = \Lambda^A = diag\{\lambda_i^A\}. \tag{5}$$

It can be shown that every $\lambda_i^A \geq 0$, as is the case for every overlap matrix.

For each atom A, we can define n_A ($n_A \leq n$) "effective atomic orbitals" $\chi_\mu^A(\vec{r})$ as linear combinations of the

"intraatomic" parts $\varphi_i^A(\vec{r})$ of the MOs as

$$\chi_\mu^A(\vec{r}) = \frac{1}{\sqrt{\lambda_\mu^A}} \sum_{i=1}^{n} U_{i\mu}^A \varphi_i^A(\vec{r}) \; ; \quad \mu = 1, 2, \ldots, n_A, \qquad (6)$$

where n_A is the number of non-zero eigenvalues λ_i^A.

One may consider rewriting the original MOs as linear combinations of the set of "effective AOs" with non-zero eigenvalue of different atoms, with the appealing result that the atomic populations calculated by the Hilbert-space analysis in the basis of the "effective AOs" are **equal** to those obtained by the 3D AIM analysis. The proof is given in Appendix A. Similarly, it can be seen that the classical Wiberg bond orders index calculated on the orthogonal basis of "effective AOs" exactly coincides with the bond orders[15] or (delocalization indices[35]).

In practice, the dimension of the "effective AO" basis has nothing to do with the dimension of the LCAO basis (if any) used in the original calculation. Equation (6) indicates that the maximum number of "effective AOs" that can be obtained is $N_{at} \times n$, which can be both less or more than the total number of the LCAO basis functions. Indeed, the proof provided in Appendix A considers up to n "effective AOs" per atom.

Moreover, experience shows that the number of "effective AOs" with significant occupation numbers on each atom is limited, and typically much smaller than the number n of the doubly occupied orbitals in the whole molecule. The remaining "effective AOs" with very small occupation numbers should have a marginal significance. Thus, one may think over expressing the molecular orbitals, to a good approximation as a linear combination of a (numerical) atomic basis set build up from a subset of the "effective AOs" (selected by an occupation number criterion), in an expansion similar to that of Eq. (A3) but with $n_A < n$. Since only a limited number (compared to the number of basis functions) of "effective AOs" typically exhibit significant occupation numbers, each molecular orbital could be expressed as a linear combination of a much reduced set of orthogonal basis functions with distinct atomic character, which may permit a much simpler analysis in chemical terms, specially when the MOs are expanded over an extended AO basis.

In order to obtain the new LCAO coefficients, one can make use of the singular value decomposition (SVD) technique to perform a pairing between the set of "effective AOs" from one side, and the set of doubly occupied MOs from the other. The procedure is described in Appendix B. In the Results and Discussion section, we will illustrate numerically how by this SVD process one can produce highly accurate atomic populations using a very much reduced set of "effective AOs." That is, the MOs can be expanded to a good approximation in terms of a minimal basis of "effective AOs."

Finally, it is worth to note that the scheme described above is not only applicable in the single-determinant closed-shell case. As noted by one of us,[28] the "effective AOs" of atom A can also be obtained from the diagonalization of the matrix $\mathbf{PS}^A$, where $\mathbf{P}$ is the LCAO density matrix and $\mathbf{S}^A$ is the atomic overlap matrix in the actual (AO or MO) basis. (In the single determinant case, if the dimension of the atomic basis applied is greater than the number of the occupied orbitals, there will be an appropriate number of strictly zero eigenvalues.)

This alternative also permits the straightforward generalization of the method both to the unrestricted single-determinant case and the correlated level, from which the $\mathbf{P}$ matrix is readily available. In that case, the SVD method can also be applied with some necessary adjustments, as indicated in Appendix B.

COMPUTATIONAL DETAILS

We have obtained the "effective AOs" in the framework of QTAIM for a series of molecules for illustrative purposes. The analysis has been performed by our program apost-3D.[36] The program includes a version of the grid-based scheme proposed by Rodríguez *et al.*[37] to integrate over atomic domains without the explicit calculation of the zero-flux surface. For the orbital plots, we have generated relatively large cubic atom-centered grids of $60 \times 60 \times 60$ points. The truncated nature of these orbitals makes rather difficult to obtain high-quality plots. An isosurface value of 0.15 has been used for all plots.

The *ab initio* calculations have been carried out with GAUSSIAN 03[38] program, employing the B3LYP density functional combined with cc-pVTZ basis set, unless otherwise indicated. For the triradical nitrene of Figure 6, we have also obtained the wavefunction at the complete active space self-consistent field (CASSCF) level of theory with an active space composed by 9 electrons and 9 orbitals. The active space includes the six orbitals of the π system of the ring, the σ orbital of the radical carbon, and the σ and π orbitals on the nitrogen atom. All calculations have been carried out at the geometrical structure of the molecules optimized at the current level of theory.

RESULTS AND DISCUSSION

Figure 1 depicts the "effective AOs" obtained for the C atom of methane. The corresponding occupation number is indicated below each orbital plot. As in methane there are only five doubly occupied MOs (in the single-determinant description), one can obtain up to five "effective AOs" for each atom. Thus, in this case there appear no weakly occupied ones on the carbon.

The first orbital can be clearly identified as a core 1*s*-type orbital with an occupation number of 2.000. The next "effective AO" is a 2*s*-type orbital, with an occupation number

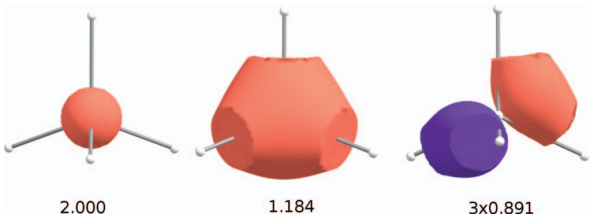

FIG. 1. "Effective AOs" for C atom in methane and their occupation numbers.

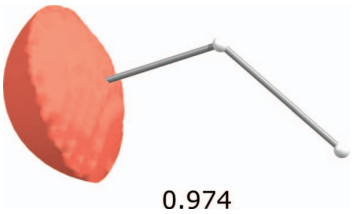

FIG. 2. "Effective AO" for H atom in methane and its occupation number.

of 1.189. The last plot of Figure 1 corresponds to one of the three degenerated (in the present context we use this term to denote orbitals with the same occupation numbers) 2*p*-type atomic orbitals.

The main feature of the "effective AOs" is that they are completely confined within the corresponding atomic domain, i.e., in this case they are strictly cut at the boundary limit of the Bader atom. We have chosen a relatively small isocontour value on the plots to highlight this feature. The occupation numbers are typically close to 2 for core orbitals or lone pairs, while hybrids that are involved in bonding (valence "effective AOs") present occupations that oscillate around 1, depending upon how the atom is polarized by the presence of the neighboring atoms. The most electronegative atoms exhibit higher occupation numbers than the less electronegative ones. For instance, for the CF_4 molecule the "effective AOs" of the C atom are strikingly similar to those of Figure 1 (only smaller because the volume of the C atom is reduced), and only the occupation numbers differ. In this case, the occupations of the 2*s*- and 2*p*-type orbitals decrease to 0.444 and 0.296, respectively.

In the case of the H atoms of methane, only the truncated *s*-type orbital shown in Figure 2 has a significant occupation. There are two additional "effective AOs" of *p*-type with much lower occupations (less than 0.03, not shown).

The polarization of the H-atoms is already included in the most populated effective AOs which (within the atomic domain) need not to be pure *s*-functions. Similarly for carbon atoms, although the number of effective AOs is the same as in a minimal *sp* basis, the form of the orbitals may slightly deviate from the ideal *s* or *p*, and reflect, therefore, the polarization effects. The truncation of the orbitals, of course, introduces another, much bigger, "polarization."

The "effective AOs" with significant occupation numbers of the C, N, and O atoms of fulminic acid (HCNO) are shown in Figure 3. Again, five hybrids with non-negligible occupation numbers are obtained for each atom, associated to their core and valence orbitals. The orbitals of all three atoms exhibit similar shapes, and the occupation numbers of analogous orbitals increase from C to O, following the increase of electronegativity. One can identify a 1*s*-type core orbital, with an occupation close to 2 in all cases. A somewhat distorted 2*s*-type orbital is also present in all atoms. Remarkably, a similar orbital was also obtained for the C atom of the acetone from a wavefunction expressed in terms of plane waves and in the framework of the "fuzzy" atoms.[31] The occupation number in the case of the oxygen atom is almost 2, as the atomic domain has no boundary on the direction opposite to the nitrogen atom. Both carbon and nitrogen exhibit two boundaries, so the corresponding hybrids are truncated from both sides. A pair of degenerate 2*p*-type orbitals on each atom are involved on the π system of this molecule. One can observe an apparent correspondence between the hybrids on carbon and nitrogen atoms. Their respective occupation numbers reflect that the atomic boundary between them is shifted towards the carbon atom. The two pairs of hybrids seem like "halves" (with different shares) of a common C–N π bond. However, it is worth to recall that they originate from two independent calculations. Similar complementarity is also observed between the two hybridized σ-type 2*p*-orbitals along the internuclear axis of nitrogen and oxygen atoms, with occupations of 0.944 and 1.091, respectively.

The occupation number of the sixth "effective AO" of the carbon atom is essentially zero (smaller than 0.01). In the oxygen atom it is 0.03, but in the case of the nitrogen atom there

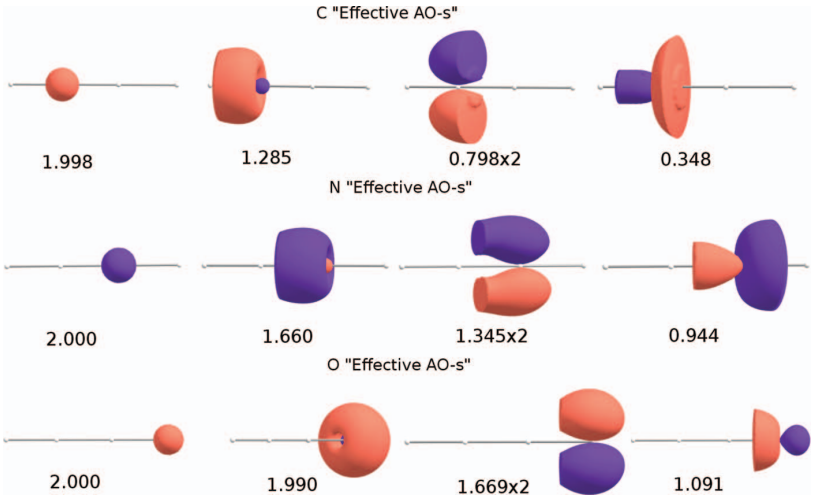

FIG. 3. "Effective AOs" and occupation numbers for the heavy atoms in fulminic acid.

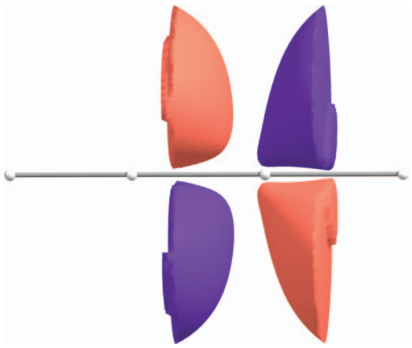

FIG. 4. Weakly occupied "effective AO" of the nitrogen atom of fulminic acid.

is a pair of degenerate "effective AOs" with a more significant occupation number of 0.133. This is a distinct manifestation of the hypervalent character of nitrogen in this molecule. In fact, Karadakov *et al.*[39] already discussed this phenomenon in this molecule with a rather involved analysis combining valence-bond and CASSCF calculations.

The shape of this weakly occupied hybrid is shown in Figure 4. It corresponds to a strongly polarized (its centroid appears displaced from the nuclear position) *d*-type orbital in a plane containing the four atoms, and with the proper orientation in order to be involved in the pi-bonding system of the molecule. Even though the occupation number is small, its participation is necessary to explain how this nitrogen atom can be involved in (formally) more than four covalent bonds (three with the carbon atom and two with the oxygen atom).

Another example of hypervalent behaviour is given by the sulfur atoms in the series $CH_3SO_xCH_3$, x = 0, 1, 2. In this case, occupation numbers of the "effective AOs" of the sulfur atoms are displayed in Figure 5.

The first observation is that the sulfur atoms exhibit in all cases five "effective AOs" with occupation number equal to 2, that correspond to the doubly occupied 1*s*, 2*s*, 2*p* inner shell. Inspection of the curve of the sulfur of dimethyl sulfide

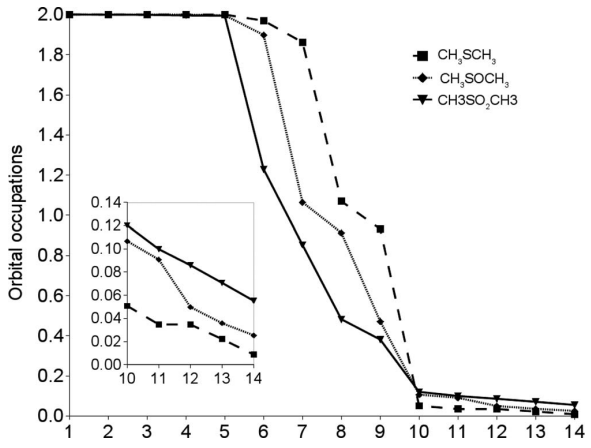

FIG. 5. Occupation numbers (in descending order) of the "effective AOs" of the sulfur atom of dimethyl sulfide, dimethyl sulfoxide, and dimethyl sulfone molecules.

(squares) reveals the presence of two more almost doubly occupied hybrids, corresponding to lone pairs. The first has a strong 3*s* character while the second one is of 3*p*-type, perpendicular to the plane formed by the sulfur and the two carbon atoms. The next two hybrids have an occupation close to 1 and are oriented towards the carbon atoms, indicating that they are involved in the sulfur-carbon σ bonds. The occupations from the ninth hybrid drop to a very small value (~0.05). Hence, the number of "effective AOs" with significant occupation number is again equal to the number of orbitals in the classical minimal basis (nine orbitals for 3rd row elements).

The curve of dimethyl sulfoxide (diamonds) indicates that there is only one lone pair. The 3*p*-type hybrid that in the case of dimethyl sulfide corresponded to a lone pair, now appears oriented towards the oxygen atom with occupation number of 0.472. Such decrease in the occupation is caused by the stronger electronegativity of the oxygen atom. The shape of the two remaining hybrids involved in the sulfur-carbon σ bonds is changed by the presence of the oxygen atom, but their occupation numbers are remarkably similar to those obtained in the case of dimethyl sulfide. In the inset of Figure 5, one can see that in this case there appear two more"effective AOs" with occupation numbers close to 0.10. These hybrids are of *d*-type and contribute to provide a slight π character to the sulfur-oxygen bonding. Similar to the case of the central N atom in fulminic acid, the participation of these *d*-orbitals is necessary to account for the formal double bond between the S and O atoms. Finally, for the sulfur atom of the dimethyl sulfone molecule there appear no doubly occupied valence hybrids. Instead, the two hybrids that are to be involved in the bonding with the oxygen atoms now exhibit occupations of 0.482 and 0.381. Moreover, there is a larger number of "effective AOs" with significant occupation numbers beyond the valence shell. The inset of the figure shows that there are essentially four hybrids of *d*-type that have a small but non-negligible contribution. The occupation number of the fifth *d*-type orbital (0.055) is very similar to that of the ninth hybrid in the case of dimethyl sulfide molecule. Such a behaviour is characteristic for the effective AOs of the hypervalent sulfur in the framework of the Hilbert space analysis, too; for an early example see Ref. 27.

For illustrative purposes, we also consider a singular nitrene triradical[40] molecule with a quartet ground state. We have obtained the "effective AOs" from the wavefunctions calculated for the quartet ground state at the CASSCF(9,9) and B3LYP levels of theory. The respective occupation numbers for selected atoms (for their numeration see Fig. 6) are gathered on Table I. The pictures of the orbitals do not differ significantly and are not shown.

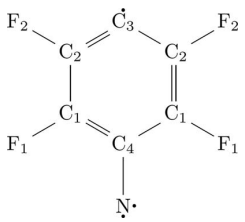

FIG. 6. Structure of the quartet fluorophenyl nitrene triradical.

TABLE I. Occupation numbers (larger than 0.01) of selected atoms of the nitrene triradical (see Fig. 6) obtained at the B3LYP and CASSCF(9,9) (in parentheses) levels of theory.

Orbital	C_1	C_3	C_4	F_1	N
1*s*	1.999 (1.999)	2.000 (2.000)	2.000 (2.000)	2.000 (2.000)	2.000 (2.000)
2*s*	1.131 (1.133)	1.295 (1.265)	1.074 (1.053)	1.998 (1.999)	1.991 (1.995)
2*p*[a]	0.916 (0.899)	0.969 (0.960)	0.862 (0.819)	1.914 (1.951)	1.144 (1.121)
2*p*[b]	0.845 (0.861)	0.865 (0.857)	0.793 (0.784)	1.942 (1.959)	0.983 (0.982)
2*p*[c]	0.342 (0.283)	0.769 (0.779)	0.584 (0.486)	1.639 (1.702)	1.349 (1.461)
d-type	0.034 (0.036)	0.039 (0.037)	0.032 (0.030)	0.061 (0.069)	0.034 (0.048)
d-type	0.030 (0.026)	0.018 (0.016)	0.029 (0.028)	0.044 (0.050)	0.032 (0.044)
d-type	0.027 (0.023)	0.013 (0.015)	0.027 (0.020)	0.027 (0.026)	0.012 (0.014)
d-type	0.025 (0.021)	0.011 (0.010)	0.017 (0.013)		

[a]Out-of-plane.
[b]In-plane.
[c]In-plane towards the center of the ring.

The occupation numbers obtained with the two methods are strikingly similar, not only between the strongly occupied "effective AOs" but also between the weakly occupied ones. The largest difference in the occupations is only about 0.10, found for a 2*p* hybrids of the N and C_6 atoms. This indicates a slight change in the polarity of the C_4–N bond going from B3LYP to CASSCF. Each of the atoms contributes to the π system with roughly one electron except for F atoms, in which the 2*p* "effective AOs" is essentially doubly occupied (1.914). The 2*p* hybrids directed towards the center of the ring exhibit different occupation numbers from one atom to another. These "effective AOs" are the ones that are involved in forming the σ-type C–F and C–N bonds, and the radical center in C_3. Thus, the occupation numbers of the "effective AOs" for C_4 and N (0.584 and 1.349) and for C_1 and F (0.342 and 1.639) exhibit a good correspondence. That close parallel between the unrestricted B3LYP and CASSCF results may be understood if one takes into account that the CASSCF method accounts for valence correlation only. The use of a method which is able to account for the (basically intra-atomic) angular correlation would perhaps make the results more complicated – but hardly lead to a different overall picture.

The picture of Fig. 6 where one unpaired electron sits on the C_3 atom and two unpaired electrons sit on the N atom has no straight correspondence with the "effective AO" analysis given in Table I. In order to locate individual electrons (of a given spin), one can obtain the "effective AOs" from the alpha and beta densities separately.[41]

TABLE II. Occupation numbers of the valence "effective AOs" of the radical centers of the nitrene triradical (see Fig. 6) obtained at the B3LYP and CASSCF(9,9) (in parentheses) levels of theory.

Orbital	C_3 alpha	C_3 beta	N alpha	N beta
2*s*	0.528 (0.506)	0.577 (0.566)	0.996 (0.997)	0.996 (0.997)
2*p*[a]	0.613 (0.651)	0.354 (0.309)	0.856 (0.854)	0.314 (0.287)
2*p*[b]	0.431 (0.419)	0.415 (0.419)	0.935 (0.952)	0.000 (0.000)
2*p*[c]	0.912 (0.936)	0.000 (0.000)	0.697 (0.723)	0.675 (0.723)

[a]Out-of-plane.
[b]In-plane.
[c]In-plane towards the center of the ring.

The respective occupation numbers for the two radical centers are collected on Table II. Note that the values of Table I are not exactly expressed as the sum of the occupation numbers of the alpha and beta contributions, as the "effective AOs" for each spin case are different. The differences in the occupation numbers between B3LYP and CASSCF are again minimal (less than 0.05). The radical center C_3 is characterized by an alpha "effective AO" with occupation close to 1, pointing outside the center of the ring. For the remaining "effective AOs" of this atom, the alpha and beta occupation numbers are similar, except for the out-of-plane hybrids forming the π system of the molecule. In that case, the occupation number of the alpha part almost doubles that of the beta. The picture of the other radical center (N atom) is very similar. The occupation numbers are in general larger than for the C_3 atom, due to the more electronegative character of nitrogen. There is a singly occupied "effective AO" of the alpha part, that lies in the molecular plane, perpendicular to the C_4–N bond. But there is no singly occupied 2*p* orbital perpendicular to the molecular plane. Instead, the occupation numbers of these alpha and beta "effective AOs" are not too different from those obtained for C_3, and as a matter of fact, for the rest of C atoms. Thus, the chemical picture that emerges from the analysis of the "effective AOs" is that there are two σ-type radical centers located on C_3 and N atoms, but there is a completely delocalized unpaired π-electron, as it is expected for such a benzonoid species. This is in agreement with the picture obtained from a local spin analysis.[42]

Finally, we illustrate how by the SVD procedure one can expand the MOs to a very good accuracy using a reduced subset of "effective AOs." We have applied the method for the alanine molecule computed with two different basis sets, namely, the 6-31++G** and the cc-pVTZ. The results are collected on Table III. We have used three different thresholds for the occupation numbers to select the subset of "effective AOs." For the cc-pVTZ basis set, up to 119 "effective AOs" have occupation number above a threshold of 0.001. This number decreases to just 71 for a threshold of 0.01, and to just 37 for a threshold of 0.1. Indeed, for this molecule the minimal basis set (1*s* for H atoms and 1*s*2*s*2*p* for the rest) includes 37 atomic orbitals. Once the "effective AOs" are selected, the SVD procedure is used to obtain the LCAO

TABLE III. Accuracy of the SVD procedure for Ala molecule with two basis sets (number of basis functions in parentheses). N_e indicates the number of effective atomic orbitals with occupation number above the threshold. $\bar{\delta}$ and δ_{max} are the average and maximum error in the atomic population values after the SVD procedure.

	6-31++G(d,p) (156)	cc-pVTZ (278)
	Occ. number >0.001	
N_e	114	119
$\bar{\delta}$	8.9×10^{-4}	8.8×10^{-4}
δ_{max}	1.9×10^{-3}	1.6×10^{-3}
	Occ. number >0.01	
N_e	69	71
$\bar{\delta}$	7.6×10^{-3}	7.3×10^{-3}
δ_{max}	1.7×10^{-2}	1.6×10^{-2}
	Occ. number >0.1	
N_e	37	37
$\bar{\delta}$	3.2×10^{-2}	3.1×10^{-2}
δ_{max}	9.8×10^{-2}	8.9×10^{-2}

coefficients of the MOs in the new numerical atomic basis. Then, the atomic populations are calculated with classical Mulliken population analysis on this basis set.

The $\bar{\delta}$ and δ_{max} on Table III account for the average and maximum error in the recomputed atomic populations, with respect to those originally obtained by the 3D-space QTAIM method. It is worth to recall that the total number of electrons is conserved after this transformation. Thus, any lack of flexibility of the numerical basis set to expand the MOs will be translated in fluctuations of the atomic populations. With a basis of 119 "effective AOs" the errors introduced are essentially within the accuracy of the numerical integration. With a basis of 71 elements, the average and maximum errors in the atomic population are 0.007 and 0.016, respectively. In case of using only the minimal basis set of 37 elements, the errors increase but still one can see that the atomic populations can be reproduced within 0.1 electrons. Interestingly, these results appear to be almost independent of the original basis set.

It is worth to note that the use of the "effective AOs" as a numerical basis set is not restricted to the framework of QTAIM of disjunct atomic domains. In fact, in the framework of "fuzzy" atoms the "effective AOs" do not strictly form an orthogonal basis set but on the other hand they can be used to construct in a systematic manner numerical basis sets of increasing accuracy for atoms (or functional groups), that could be used in fully numerical[43] *ab initio* calculations. Of course, one should expect that in the case of correlated methods the size of the numerical basis set should be larger, for a better description of the virtual orbitals.

CONCLUSIONS

The "effective AOs" have been realized in the framework of Bader's QTAIM. This formalism can be used to retrieve from any type of calculation a proper set of orthogonal atomic basis functions. They form an orthonormalized set of numerical atomic orbitals, with occupation numbers that sum up to the respective QTAIM atomic populations. Importantly, only a limited number of "effective AOs" exhibit significant occupation numbers, i.e., these atomic hybrids closely mimic the core and valence shells of the atom. In the case of hypervalent atoms, there appear additional hybrids with small but non-negligible occupation numbers.

We have shown that the MOs can be exactly expressed as a linear combination of this orthonormalized set of numerical atomic orbitals. Moreover, the Mulliken population analysis carried out on the basis of "effective AOs" exactly reproduces the original QTAIM atomic populations of the atoms. Approximate expansion of the MOs over a much reduced set of orthogonal atomic basis functions can also be accomplished to a very good accuracy with a SVD procedure.

Thus, this shows that there is nothing fundamentally inappropriate with a Hilbert-space based population analysis. The flaws of the classical Mulliken populations are rooted in the use of unsuitable atomic basis functions, not in its mathematical framework.[44]

ACKNOWLEDGMENTS

Financial help has been furnished by the Spanish MICINN Project No. CTQ2011-23441/BQU and Acción Complementaria del MCI (PCI2006-A7-0631). Financial support from MICINN and the FEDER fund (European Fund for Regional Development) was also provided by grant UNGI08-4E-003. Financial support from the Generalitat de Catalunya (SGR528 and Xarxa de Referncia en Qumica Terica i Computacional) is also acknowledged. E.R-C. acknowledges support from the Spanish FPU program (Grant No. AP2008-01231).

I.M. acknowledges partial financial support of the Hungarian Scientific Research Fund (Grant No. OTKA 71816).

APPENDIX A: EQUIVALENCE OF HILBERT-SPACE ANALYSIS IN THE BASIS OF EFFECTIVE AOs AND QTAIM ANALYSIS

By virtue of the definition (2), orbitals $\chi_\mu^A(\vec{r})$ differ from zero only in the atomic domain of atom A. They are orthonormalized

$$\begin{aligned}\langle \chi_\mu^A | \chi_\nu^A \rangle &= \left\langle \frac{1}{\sqrt{\lambda_\mu^A}} \sum_{i=1}^{n} U_{i\mu}^A \varphi_i^A \middle| \frac{1}{\sqrt{\lambda_\nu^A}} \sum_{j=1}^{n} U_{j\nu}^A \varphi_j^A \right\rangle \\ &= \frac{1}{\sqrt{\lambda_\mu^A \lambda_\nu^A}} \sum_{i,j=1}^{n} (\mathbf{U}^{A\dagger})_{\mu i} Q_{ij}^A U_{j\nu}^A \\ &= \frac{1}{\sqrt{\lambda_\mu^A \lambda_\nu^A}} \lambda_\mu^A \delta_{\mu\nu} = \delta_{\mu\nu} \end{aligned} \tag{A1}$$

as a consequence of the eigenvalue equation (5). In fact, orbitals χ_μ^A represent the functions obtained by performing Löwdin's "canonic" orthogonalization of the functions φ_i^A. (Not to be confused with the usual Löwdin-orthogonalization performed by using matrix $\mathbf{S}^{-1/2}$.)

Relationship (6) can be trivially inverted, and one gets

$$\varphi_j^A = \sum_{\mu=1}^{n_A} U_{j\mu}^{A*} \sqrt{\lambda_\mu^A} \chi_\mu^A. \tag{A2}$$

Owing to this result and Eq. (4), the MOs can be written as linear combinations of the "effective AOs" of different atoms

$$\varphi_i = \sum_{A}^{N_{at}} \sum_{\mu=1}^{n_A} U_{i\mu}^{A*} \sqrt{\lambda_\mu^A} \chi_\mu^A. \tag{A3}$$

Thus, our analysis of the behaviour of the molecular orbitals in the different atomic domains resulted in a special LCAO expansion of the molecular orbitals. One can also introduce a continuous numbering of the basis orbitals, and write

$$\varphi_i = \sum_{\nu=1}^{m} C_{\nu i} \chi_\nu, \tag{A4}$$

where the overall dimension of the "effective AO" basis is

$$m = \sum_{A=1}^{N_{at}} n_A, \tag{A5}$$

and the orbitals χ_ν with $\nu \le n_1$ are attributed to atom with number $A = 1$, those with $n_1 + 1 \le \nu \le n_1 + n_2$ to atom $A = 2$, and so on, For the LCAO coefficients $C_{\nu i}$, one obviously has

$$C_{\nu i} = U_{i\mu}^{A*} \sqrt{\lambda_\mu^A}. \tag{A6}$$

If orbital χ_ν belongs to atom A, then subscripts μ and ν are trivially related as

$$\mu = \nu - \sum_{B=1}^{A-1} n_B. \tag{A7}$$

Owing to the orthogonality relationship (A1) and the disjunct character of the atomic basins, the orbitals χ_ν, $\nu = 1, 2, \ldots, m$ form an orthonormalized basis. Therefore, there are no overlap populations, and Mulliken's net and gross populations coincide. Thus, the Hilbert space or LCAO population of atom A in terms of the "effective atomic orbitals" is given by the sum of the diagonal density matrix elements for the orbitals belonging to that atom

$$\begin{aligned} Q_A^{LCAO} &= \sum_{\nu \in A} D_{\nu\nu} = 2 \sum_{i=1}^{n} \sum_{\nu \in A} |C_{\nu i}|^2 = 2 \sum_{i=1}^{n} \sum_{\mu=1}^{n_A} \left| U_{i\mu}^{A*} \sqrt{\lambda_\mu} \right|^2 \\ &= 2 \sum_{i=1}^{n} \sum_{\mu=1}^{n_A} U_{i\mu}^{A} U_{i\mu}^{A*} \lambda_\mu = \sum_{i=1}^{n} \sum_{\mu,\nu=1}^{n_A} U_{i\mu}^{A} U_{i\nu}^{A*} \lambda_\mu \delta_{\mu\nu} \\ &= \sum_{i=1}^{n} \sum_{\mu,\nu=1}^{n} U_{i\mu}^{A} \Lambda_{\mu\nu} (\mathbf{U}^{A\dagger})_{\nu i} \\ &= 2 \sum_{i=1}^{n} (\mathbf{U}^A \Lambda^A \mathbf{U}^{A\dagger})_{ii} = 2 \sum_{i=1}^{n} Q_{ii}^A, \end{aligned} \tag{A8}$$

where the inverse of Eq. (5) has been utilized. (The summation limit for μ, ν was increased from n_A to n because that meant only adding terms containing factors $\lambda_\mu = 0$.)

The AIM population of atom A is given by

$$\begin{aligned} Q_A^{AIM} &= \int_{\Omega_A} \rho(\vec{r}) dv = 2 \int_{\Omega_A} \sum_{i=1}^{n} |\varphi_i(\vec{r})|^2 dv \\ &= 2 \sum_{i=1}^{n} \int_{\Omega_A} |\varphi_i(\vec{r})|^2 dv = 2 \sum_{i=1}^{n} Q_{ii}^A. \end{aligned} \tag{A9}$$

Comparing Eqs. (A8) and (A9), we see that

$$Q_A^{LCAO} = Q_A^{AIM}, \tag{A10}$$

i.e., the atomic population calculated by the Hilbert-space analysis in the basis of the "effective AOs" is equal to that obtained by the 3D AIM analysis.

APPENDIX B: APPROXIMATE EXPANSION OF THE MOs IN THE BASIS OF "EFFECTIVE AOs"

Let us consider those "effective AOs" χ_μ^A, $A = 1, 2, \ldots N_{at}$ which meet some criterion $\lambda_\mu^A \ge t \ge 0$. Let their effective number be n_{eff}. In order to get an (approximate) expansion of the MOs, we should take enough "effective AOs," so it must be $n_{eff} \ge n$.

Now we build the rectangular $n_{eff} \times n$ matrix $\mathbf{Z}$, with elements

$$Z_{\mu i} = \langle \chi_\mu | \varphi_i \rangle. \tag{B1}$$

With the SVD, the rectangular matrix is transformed as

$$\mathbf{U}^\dagger \mathbf{Z} \mathbf{V} = \Xi, \tag{B2}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices of dimension $n_{eff} \times n_{eff}$ and $n \times n$, respectively, and Ξ is a rectangular diagonal matrix containing the singular values ξ_i of $\mathbf{Z}$. From the definition of matrix $\mathbf{Z}$ and using Eq. (B2), one can write

$$\sum_{\mu}^{n_{eff}} \sum_{j}^{n} U_{\mu i}^* \langle \chi_\mu | \varphi_j \rangle V_{ji} = \xi_i, \tag{B3}$$

that is, the singular value ξ_i is the overlap between the function

$$\psi_i = \sum_{\mu=1}^{n_{eff}} U_{\mu i} \chi_\mu, \tag{B4}$$

which is a linear combination of the "effective AOs" and

$$\varphi_i' = \sum_{\mu=1}^{n} V_{ji} \varphi_j, \tag{B5}$$

representing a molecular orbital after performing a unitary transformation with the matrix $\mathbf{V}$.

If $\xi_i = 1$, the two functions have an overlap equal one, ψ_i and φ_i' are essentially ("almost everywhere") equal to each other, and one can write

$$\varphi_i' = \sum_{\mu}^{n_{eff}} U_{\mu i} \chi_\mu. \tag{B6}$$

That is, the columns of the unitary matrix $\mathbf{U}$ contain the LCAO coefficients of each rotated MO in the orthogonal basis of "effective AOs."

If ξ_i is close to, but not exactly equal one, Eq. (B6) represents an approximation to the (rotated) molecular orbital. Under these circumstances, Eq. (A10) is no longer strictly fulfilled but, since these approximated MOs form an orthonormalized set, the number of electrons is conserved. (The "effective AOs" χ_μ are orthonormalized and matrix **U** is unitary, so we have n orthonormalized approximate MOs φ_i', so the closed shell determinant wavefunction constructed by their use carries exactly as many electrons as the original wavefunction containing the orbitals φ_i.)

Finally, the SVD method can also be applied in practice beyond doubly occupied orbitals. First of all, in order to reduce the dimensionality of the problem, it is worth to consider the natural orbital representation, ψ_i^{nat}, and thus perform the pairing between the set of "effective AOs" on one side and a reduced number of natural orbitals, n_{occ} (chosen again by an occupation number criterion). After the SVD process, each rotated natural orbital is identified with a linear combination of "effective AOs," as in the previous case

$$\sum_j^{n_{occ}} V_{ji}\psi_j^{nat} \cong \sum_\mu^{n_{eff}} U_{\mu i}\chi_\mu. \tag{B7}$$

However, now it is more convenient to have the connection directly with the original set of natural orbitals, where the density matrix is diagonal. Multiplying Eq. (B7) by the matrix element $(\mathbf{V}^\dagger)_{ik} = V_{ki}^*$, and summing up over i, one gets

$$\psi_k^{nat} \cong \sum_\mu^{n_{eff}}\sum_i^{n_{occ}} U_{\mu i}(\mathbf{V}^\dagger)_{ik}\chi_\mu = \sum_\mu^{n_{eff}} W_{\mu k}\chi_\mu, \tag{B8}$$

where now the columns of matrix $\mathbf{W} = \mathbf{U}\mathbf{V}^\dagger$ gather the (approximate) LCAO coefficients of the original subset of natural molecular orbitals over the "effective AO" basis.

[1] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
[2] P.-O. Löwdin, Adv. Quantum Chem. **5**, 185 (1970).
[3] F. L. Hirshfeld, Theor. Chim. Acta **44**, 129 (1977).
[4] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, Oxford, 1990).
[5] I. Mayer and P. Salvador, Chem. Phys. Lett. **383**, 368 (2004).
[6] T. C. Lillestolen and R. J. Wheatley, Chem. Commun. **45**, 5909 (2008).
[7] P. Bultinck, C. Van Alsenoy, P. W. Ayers, and R. Carbó-Dorca, J. Chem. Phys. **126**, 144111 (2007).
[8] T. C. Lillestolen and R. J. Wheatley, Chem. Commun. **2008**, 5909.
[9] P. Bultinck, D. L. Cooper, and D. V. Neck, Phys. Chem. Chem. Phys. **11**, 3424 (2009).
[10] T. C. Lillestolen and R. J. Wheatley, J. Chem. Phys. **131**, 144101 (2009).
[11] I. Mayer, J. Comput. Chem. **28**, 204 (2007).
[12] J. Baker, Theor. Chim. Acta **68**, 221 (1985).
[13] E. R. Davidson, J. Chem. Phys. **46**, 3320 (1967).
[14] R. Heinzmann and R. Ahlrichs, Theor. Chem. Acc. **42**, 33 (1976).
[15] J. G. Ángyán, M. Loos, and I. Mayer, J. Phys. Chem. **98**, 5244 (1994).
[16] E. Ramos-Cordoba, E. Matito, P. Salvador, and I. Mayer, Phys. Chem. Chem. Phys. **14**, 15291 (2012).
[17] P. Salvador and I. Mayer, J. Phys. Chem. **120**, 5046 (2004).
[18] M. A. Blanco, A. Martín Pendás, and E. Francisco, J. Chem. Theory Comput. **1**, 1096 (2005).
[19] E. Francisco, A. M. Pendás, and M. A. Blanco, J. Chem. Theory Comput. **2**, 90 (2006).
[20] E. Ramos-Cordoba, E. Matito, I. Mayer, and P. Salvador, J. Chem. Theory Comput. **8**, 1270 (2012).
[21] M. D. Pedro Salvador and I. Mayer, J. Chem. Phys. **115**, 1153 (2001).
[22] D. R. Alcoba, L. Lain, A. Torre, and R. C. Bochicchio, Chem. Phys. Lett. **407**, 379 (2005).
[23] L. Li and R. G. Parr, J. Chem. Phys. **84**, 1704 (1986).
[24] D. R. Alcoba, L. Lain, A. Torre, and R. C. Bochicchio, J. Chem. Phys. **123**, 144113 (2005).
[25] D. Vanfleteren, D. V. Neck, P. Bultinck, P. W. Ayers, and M. Waroquier, J. Chem. Phys. **133**, 231103 (2010).
[26] R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[27] I. Mayer, Chem. Phys. Lett. **242**, 499 (1995).
[28] I. Mayer, J. Phys. Chem. **100**, 6249 (1996).
[29] I. Mayer, Can. J. Chem. **74**, 939 (1996).
[30] I. Mayer and P. Salvador, J. Chem. Phys. **130**, 234106 (2009).
[31] I. Mayer, I. Bakó, and A. Stirling, J. Phys. Chem. A **115**, 12733 (2011).
[32] R. Ponec, J. Math. Chem. **21**, 323 (1997).
[33] D. Tiana, E. Francisco, M. A. Blanco, P. Macchi, A. Sironi, and A. Martin Pendás, Phys. Chem. Chem. Phys. **13**, 5068 (2011).
[34] J. Cioslowski, Int. J. Quantum Chem. **38**, 015 (1990).
[35] X. Fradera, M. A. Austen, and R. F. Bader, J. Phys. Chem. A **103**, 304 (1999).
[36] P. Salvador and E. Ramos-Cordoba, APOST-3D, University of Girona, Spain, 2011.
[37] J. I. Rodríguez *et al.*, J. Comput. Chem. **30**, 1082–1092 (2009).
[38] M. J. Frisch *et al.*, GAUSSIAN 03, Revision C.02, Gaussian, Inc., Pittsburgh, PA, 2003.
[39] P. Karadakov, D. Cooper, T. Thorsteinsson, and J. Gerratt, "Modern Valence-Bond Description of the Mechanisms of Six-Electron Pericyclic Reactions," in *Quantum Systems in Chemistry and Physics. Volume 1: Basic Problems and Models Systems*, edited by A. Hernández-Laguna, J. Maruani, R. McWeeny and S. Wilson (Kluwer, Dordrecht, 2000) pp. 327–344.
[40] W. Sander, D. Grote, S. Kossmann, and F. Neese, J. Am. Chem. Soc. **130**, 4396 (2008).
[41] R. Ponec, E. Ramos-Cordoba, and P. Salvador, J. Phys. Chem. A **117**, 1975 (2013).
[42] E. Ramos-Cordoba and P. Salvador, "Local spin analysis of polyradicals" (unpublished).
[43] M. Zakharov, "Performance of numerical atom-centered basis sets in the ground-state correlated calculations of noncovalent interactions: Water and methane dimer cases," Int. J. Quantum Chem. (published online).
[44] I. Mayer, Chem. Phys. Lett. **97**, 270 (1983).

6.2 Oxidation states from wave function analysis

As submitted for publication to Chem. Commun.

Ramos-Cordoba, E.; Salvador, P. “Oxidation States fromWave Function Analysis”. *Chemical Communications*. Submitted, 2014.

http://pubs.rsc.org/en/journals/journalissues/cc#!recentarticles&all

Abstract

We introduce a simple and general scheme to retrieve oxidation states from first principle calculations. The method relies on the so-called spin-resolved effective atomic orbitals (eff-AO-s). For practical applications a strategy based on the partitioning of the molecular system into fragment/ligands prior to the eff-AO analysis is proposed.

Chapter 7

The elusive $[Cu_3S_2]^{3+}$ core: a case of study

7.1 Bonding quandary in the $[Cu_3S_2]^{3+}$ core: insights from the analysis of domain averaged Fermi holes and the local spin

Ponec, R.; Ramos-Cordoba, E.; Salvador, P. “Bonding Quandary in the [Cu3S2]3+ Core: Insights from the Analysis of Domain Averaged Fermi Holes and the Local Spin”. *Journal of Physical Chemistry A*. 2013, 117, 1975–1982

http://dx.doi.org/10.1021/jp309295r

http://pubs.acs.org/doi/abs/10.1021/jp309295r

Received: September 19, 2012

Revised: February 4, 2013

Published: February 7, 2013

Abstract

The electronic structure of the trinuclear symmetric complex $[(tmedaCu)_3S_2]^{3+}$, whose Cu_3S_2 core represents a model of the active site of metalloenzymes involved in biological processes, has been in recent years the subject of vigorous debate. The complex exists as an open-shell triplet, and discussions concerned the question whether there is a direct S–S bond in the $[Cu_3S_2]^{3+}$ core, whose answer is closely related to the problem of the formal oxidation state of Cu atoms. In order to contribute to the elucidation of the serious differences in the conclusions of earlier studies, we report in this study the detailed comprehensive analysis of the electronic structure of the $[Cu_3S_2]^{3+}$ core using the methodologies that are specifically designed to address three particular aspects of the bonding in the core of the above complex, namely, the presence and/or absence of direct S–S bond, the existence and the nature of spin–spin interactions among the atoms in the core, and the formal oxidation state of Cu atoms in the core. Using such a combined approach, it was possible to conclude that the picture of bonding consistently indicates the existence of a weak direct two-center–three-electron (2c–3e) S–S bond, but at the same time, the observed lack of any significant local spin in the core of the complex is at odds with the suggested existence of antiferromagnetic coupling among the Cu and S atoms, so that the peculiarities of the bonding in the complex seem to be due to extensive delocalization of the unpaired spin in the $[Cu_3S_2]^{3+}$ core. Finally, a scrutiny of the effective atomic hybrids and their occupations points to a predominant formal CuII oxidation state, with a weak contribution of partial CuI character induced mainly by the partial flow of electrons from S to Cu atoms and high delocalization of the unpaired spin in the $[Cu_3S_2]^{3+}$ core.

Chapter 8

Computer implementation: APOST-3D program

The new methods developed in this thesis have been implemented in a new program called APOST-3D.[89] The origins of the new software were the existing codes AFUZZY[90] and ENPART.[91] The former was devised to compute charges, bond orders, overlap populations and valences alongside with effective atomic orbitals in the framework of the "fuzzy" atoms. The latter was designed to compute energy partitioning for fuzzy atoms for either density functional theory and wave function methods. Both programs were written in our laboratory by Pedro Salvador and have been recently united to release APOST-3D. The abovementioned AFUZZY and ENPART, as well as a number of other programs produced by Prof. I. Mayer have been made available to the scientific community on the webpage http://occam.chemres.hu/programs/. The present APOST-3D code will also be made available in the near future, along with a short manual.

The main feature of APOST-3D is that it performs fast numerical integrations of one- and two-electron functions that can be easily split into atomic or diatomic domains. Nowadays it can perform integration over atomic domains using different atomic definitions such as Hirshfeld, Hirshfeld-Iterative, Becke, Becke-rho, TFVC and QTAIM. The program can be used to compute atomic populations, bond orders or delocalization indexes, free-valences, molecular en-

ergy partitions, local spins, effective atomic orbitals, effective oxidation states, conceptual DFT indexes, among other chemical properties.

The features of the program that have been implemented during the course of this thesis are the following

Atom in Molecule Definitions: We have implemented a version of the grid-based scheme proposed by Rodriguez[92] to integrate over QTAIM's atomic domains without the explicit calculation of the zero-flux surface. The new TFVC described has also been included in the software. We have also added a proper mapping between the atomic overlap matrices used in the framework of 3D-space analysis and their Hilbert-space analogues, such as Mulliken and Löwdin. In this sense, we are able to use the same algorithms for both Hilbert-space and 3D-space analysis.

Local Spin Analysis: All the decompositions for the local spin that have appeared in the literature since the first one proposed by Clark and Davidson are implemented in APOST-3D. By default the program computes the decomposition described in Chapters 6-7 as it is the only one which fulfills all the physical requirements imposed to date. To perform the local spin analysis for correlated wave functions the 2-RDM is needed and most of the commercial *ab initio* packages do not generate these matrices. To this end we have made use of the DMN program by Matito and Feixas. Our algorithm to perform the local spin analysis reads and processes the 2-RDM elements on the fly since the 2-RDM does not need to be saved in memory.

Effective Atomic Orbitals: A spin resolved version of the eff-AOs has been implemented. Morever, taking advantage of the QTAIM algorithm described above we have included the possibility to generate Cube files to plot the eff-AOs generated under the QTAIM framework. We have included a singular value decomposition approach to expand (approximately) the MOs on a reduced set of eff-AOs.

Oxidation States: We have implemented the algorithm to compute effective oxidation states (EOS) and the confidence index of the assignation. Moreover, in order to apply the EOS for large ligands, we have included the possibility to perform the eff-AOs for fragments instead of single atoms.

Chapter 9

Results and discussion

Hereinafter the main results from Chapters 3-7 will be summarized. For the sake of simplicity we have divided this Chapter in five main parts. The first part deals with the new atom in molecule definition presented in Chapter 3. In the second part the new formulation of the local spin is presented and its main properties are outlined (Chapter 4). In the third part we summarize the applications of the local spin analysis. From one side the relation between the local spin and the chemical bond is established, and from the other side we discuss the ability of the local spin analysis to characterize and quantify the polyradical character (Chapter 5). In the fourth part we collect the main achievements concerning the derivation of oxidation states from wave function analysis, relying on the information provided by the eff-AO analysis of atoms and molecular fragments. In the last part we summarize the results of the application of the local spin, eff-AO and DAFH methodologies to a challenging molecular system, namely the $[Cu_3S_2]^{3+}$ core.

9.1 On the definition of a new "fuzzy" atom scheme

Atoms are the building blocks of chemistry. Even when dealing with quantum mechanics, where the atom is not an observable, theoretical and computational chemists have the need to invoke the classical idea of atom in a molecule in

order to establish bridges between quantum and classical chemical concepts.

Several atom in molecules definitions have been reviewed in the introduction of this thesis. Bader's QTAIM is one of the most used but it is also one of the computationally most expensive. Fuzzy-atom partitions have emerged as a computationally cheaper alternative to QTAIM. The combination of Becke's atoms and the information provided by the topology of the density led to the Becke-ρ atomic definition. This atom definition provides results quantitatively similar to those obtained with Bader's topological atoms. However, the original Becke atom definition has an inherent limitation that affects atom pairs with very different size and electronegativity. In particular, it can not accommodate shifts in the interatomic boundaries between atoms whose relative size exceeds a value of 2.41.(It is important to recall here that the limitation was already noted by Becke in his original manuscript but had passed unperceived).

Chapter 3 reviews the fuzzy atomic Voronoi cells introduced by Becke, and the Becke-ρ scheme. The limitation of Becke's scheme is explained in detail and a modification of the original formulation is proposed to prevent the above mentioned limitation. The proposed modification (eq. (9.1)) affected the transformation of the atomic pair coordinate μ_{AB}, needed to account for different atomic sizes in heteronuclear systems. The new transformed coordinate reads as follows

$$\nu'_{AB} = \frac{1 + \mu_{AB} - \chi_{AB}(1 - \mu_{AB})}{1 + \mu_{AB} + \chi_{AB}(1 - \mu_{AB})}, \tag{9.1}$$

where χ_{AB} accounts for the relative size of atoms A and B.

The new transformed coordinate fulfills all the original criteria of Becke's formulation and it is monotonic for any value of χ_{AB}. In the implementation of the new scheme, called Topological Fuzzy Voronoi Cells (TFVC), we have chosen a stiffness parameter $k = 4$. Parameter k controls the shape of the atom pair cutoff profile. Interestingly, a modest grid of 40 radial and 146 angular points per atoms is needed to achieve a total number of electrons with an error of ca. 10^{-3}. For the sake of comparison recent implementations of QTAIM integrations use up to 0.87 million grid points per atom. Moreover, TFVC scales linearly with the total number of grid points.

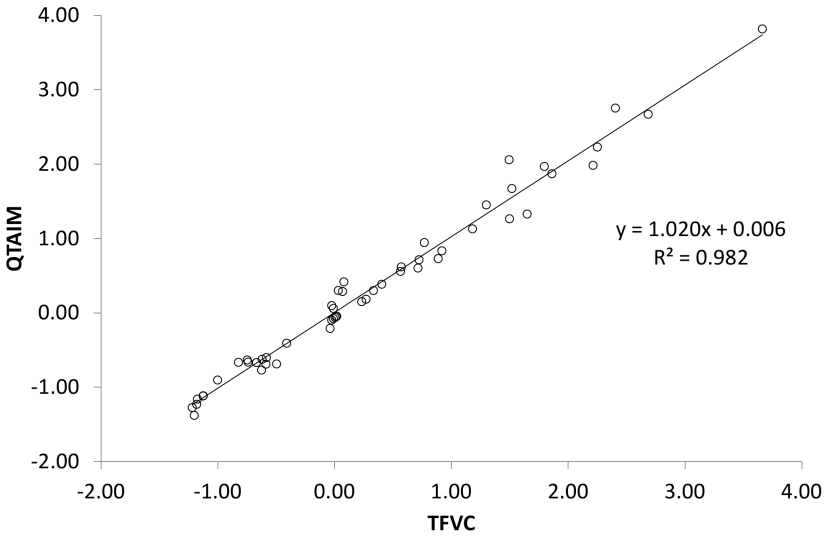

Figure 9.1: QTAIM vs TFVC atomic charges for the molecular set described in Chapter 3.

Contrary to the Becke-ρ scheme, TFVC fully avoids the use of a fixed set of atomic radii. We introduce an additional criterion, so that a given pair of atoms are not considered as neighbors if their middle point is closer to a third nucleus or a NNA of the molecular system. An important feature of TFVC is that a NNA can be either ignored or incorporated in the analysis, by defining a new Voronoi cell.

Figures (9.1) and (9.2) display the correlation between the partial atomic charges and DIs of a set of molecules (essentially a series of hydrides) at the B3LYP/cc-pVTZ level. As indicated by the R^2 and the slope values the agreement between TFVC and QTAIM charges and DIs is very good. The mean unsigned error between TFCV and QTAIM is 0.12 for charges and 0.06 for DI's.

A clear difference between QTAIM and TFVC schemes is that in the latter the cell boundaries are described by planes, while in the former are typically curved. This difference does not seem to affect the agreement between the two atomic definitions in most cases. For instance, in the case of H_2O (see fig. (9.3)) the curvature of the zero-flux surface is very pronounced so the shape of the QTAIM and TFVC atoms is quite different. Yet, the integrated electron

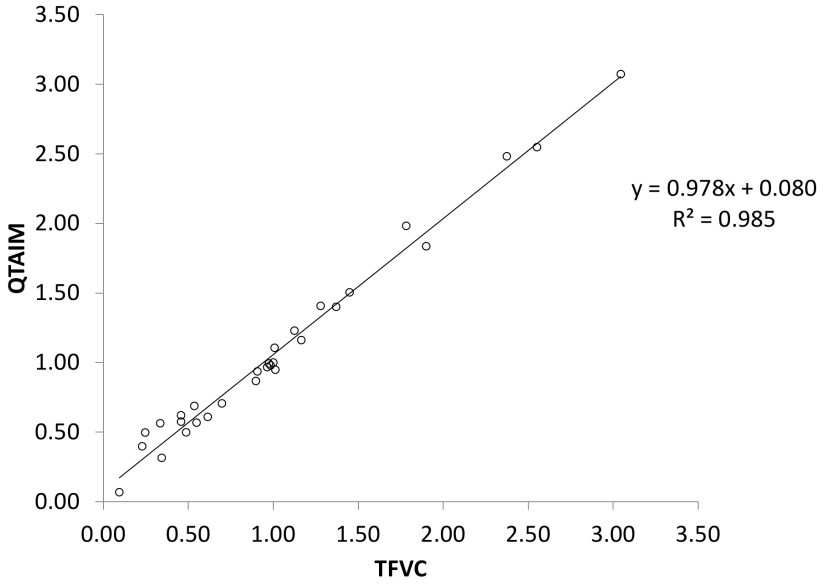

Figure 9.2: QTAIM vs TFVC delocalization indices for the molecular set described in Chapter 3.

populations are strikingly close. The same occurs with hydrogen fluoride. On the other hand, in H_2S (see fig. (9.4)) the shape of the QTAIM and TFVC boundaries are quite alike but the differences in the integrated quantities are significant (almost 1 e for the S atom). To sum up, TFVC and QTAIM reach its greater divergence when electron density is high and flat in the interatomic region.

TFVC have also been tested for the decomposition of the Hartree-Fock molecular energy. It has been compared against QTAIM and the original Becke-ρ scheme. The value of the Laplacian of the density integrated over atomic domains (an indicator of the fulfillment of the zero flux condition within QTAIM) is ca. 10^{-2} for TFCV, ca. 10^{-1} for Becke-ρ and ca. 10^{-4} for typical numerical implementations of QTAIM. The differences of the integrated Laplacian between the methods does not seem to influence too much the values of the energy contributions. In general, the TFVC integrated energy contributions are much closer to QTAIM than the Becke-ρ ones.

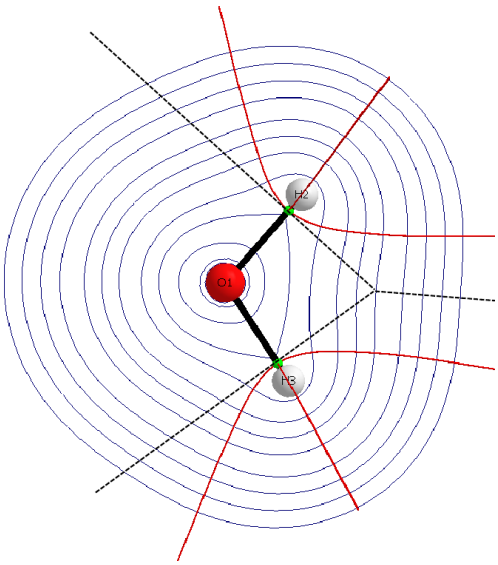

Figure 9.3: Voronoi cell boundary (dashed line) and interatomic surface paths (red lines) for H_2O in the molecular plane.

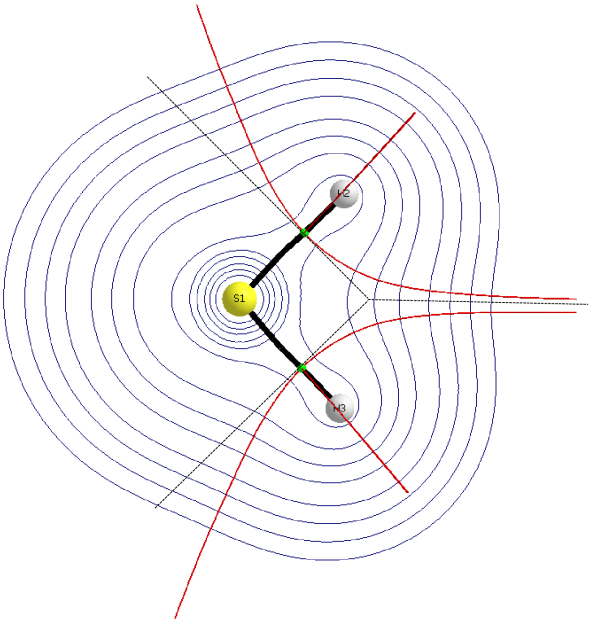

Figure 9.4: Voronoi cell boundary (dashed line) and interatomic surface paths (red lines) for H_2S in the molecular plane.

9.2 On the proper definition of local spin

The spin density is usually the most used tool to characterize the spin properties of molecular systems. However, in the case of singlet diradicals, antiferromagnets or dissociation processes of molecules the spin density is not sufficient to describe the physical situation as $\rho^s(\vec{r}) = 0$ for any pure singlet. A proper spin description of these systems can only be done with multideterminant wave functions or by broken symmetry approaches.

An alternative to the spin density is the local spin analysis. These procedures consists on the decomposition of $\langle \hat{S}^2 \rangle$ as a sum of atomic and diatomic contributions:

$$\langle \hat{S}^2 \rangle = \sum_A \langle \hat{S}^2 \rangle_A + \sum_{A,B \neq A} \langle \hat{S}^2 \rangle_{AB}. \tag{9.2}$$

Several formulations of the local spin have been proposed to date. Because there are in principle infinite ways to decompose $\langle \hat{S}^2 \rangle$ into atomic and diatomic terms, a number of physical requirements have been introduced, namely
(i) One should obtain no spins for covalent systems described by a closed-shell RHF wave function using doubly filled orbitals.

(ii) If the wave function is properly dissociating, the asymptotic values of the atomic spins obtained for the atoms at large distances should coincide with the corresponding values of the free atoms.

(iii) The formula used for a general wave function should reduce to that used in the single-determinant case if applied to single-determinant wave functions.

(iv) In an open-shell system the overall $\langle \hat{S}^2 \rangle$ does not depend on the actual $\hat{S}_z$ projection of the electronic state considered (m_S value), so one may request to have $\langle \hat{S}^2 \rangle$ components that do not depend on m_S either.

We have shown that these requirements are not sufficient to define a unique formulation of the local spin. Actually, there is a *continuum* of formulations

that fulfill the requirements above mentioned. To see that, let us write the expectation value of $\hat{S}^2$ for an N electron system as

$$\begin{aligned}\langle \hat{S}^2 \rangle &= \frac{3}{4}\int \rho(\vec{r}_1;\vec{r}_1)d\vec{r}_1 - \frac{1}{4}\int\int \rho_2(\vec{r}_1,\vec{r}_2;\vec{r}_1,\vec{r}_2)d\vec{r}_1 d\vec{r}_2 \\ &\quad -\frac{1}{2}\int\int \rho_2(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1)d\vec{r}_1 d\vec{r}_2. \end{aligned} \tag{9.3}$$

Expressing the 2-RDM in terms of the 1-RDM, the spin density, and the cumulant or the 2-RDM one obtains

$$\begin{aligned}\langle \hat{S}^2 \rangle &= \frac{3}{4}\int \rho(\vec{r}_1)d\vec{r}_1 - \frac{3}{8}\iint \rho(\vec{r}_1;\vec{r}_2)\rho(\vec{r}_2;\vec{r}_1)d\vec{r}_1 d\vec{r}_2 \\ &\quad +\frac{1}{8}\iint \rho^s(\vec{r}_1;\vec{r}_2)\rho^s(\vec{r}_2;\vec{r}_1)d\vec{r}_1 d\vec{r}_2 + \frac{1}{4}\int \rho^s(\vec{r}_1)d\vec{r}_1 \int \rho^s(\vec{r}_2)d\vec{r}_2 \\ &\quad -\frac{1}{4}\iint \Gamma(\vec{r}_1,\vec{r}_2)d\vec{r}_1 d\vec{r}_2 - \frac{1}{2}\iint \Gamma(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1)d\vec{r}_1 d\vec{r}_2. \end{aligned} \tag{9.4}$$

The spin density and the cumulants vanish for single-determinant restricted wave functions. However, the first and second terms on the r.h.s. of eq. (9.4) can be decomposed into nonzero one-center and one- and two-center contributions respectively, thus not satisfying condition i).

By using the definition of the effectively unpaired electron density $u(\vec{r})$, eq. (9.4) can be transformed to

$$\begin{aligned}\langle \hat{S}^2 \rangle &= \frac{3}{8}\int u(\vec{r}_1)d\vec{r}_1 - \frac{1}{4}\iint \left[\Gamma(\vec{r}_1,\vec{r}_2) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_2)\rho^s(\vec{r}_2;\vec{r}_1)\right] d\vec{r}_1 d\vec{r}_2 \\ &\quad -\frac{1}{2}\iint \left[\Gamma(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_1)\rho^s(\vec{r}_2;\vec{r}_2)\right] d\vec{r}_1 d\vec{r}_2. \end{aligned} \tag{9.5}$$

All the terms of the r.h.s. of eq. (9.5) vanish for single-determinant restricted wave functions. Moreover, since $u(\vec{r}_1)$ and

$$\Gamma(\vec{r}_1,\vec{r}_2;\vec{r}'_1,\vec{r}'_2) - \frac{1}{2}\rho^s(\vec{r}_1;\vec{r}'_2)\rho^s(\vec{r}_2;\vec{r}'_1)$$

are spin-independent quantities, eq. (9.5) is the appropriate starting point to

derive a decomposition fulfilling conditions *i)-iv)*. Using the relationship

$$\int \left[\Gamma(\vec{r}_1,\vec{r}_2) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_2)\rho^s(\vec{r}_2;\vec{r}_1)\right] d\vec{r}_2 = -\tfrac{1}{2}u(\vec{r}_1) \tag{9.6}$$

the first two terms of the r.h.s. of eq. (9.5) can be modulated in terms of a parameter a and obtain a general expression like

$$\begin{aligned}\langle \hat{S}^2\rangle &= a\int u(\vec{r}_1)d\vec{r}_1 - (1-2a)\iint \left[\Gamma(\vec{r}_1,\vec{r}_2) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_2)\rho^s(\vec{r}_2;\vec{r}_1)\right] d\vec{r}_1 d\vec{r}_2 \\ &\quad -\frac{1}{2}\iint \left[\Gamma(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_1)\rho^s(\vec{r}_2;\vec{r}_2)\right] d\vec{r}_1 d\vec{r}_2. \end{aligned}\tag{9.7}$$

Within the 3D-space analysis the atomic and diatomic spin components from eq. (9.7) can be obtained by introducing a non-negative weight function w_A, with

$$\begin{aligned}\langle \hat{S}^2\rangle_A &= a\int w_A(\vec{r}_1)u(\vec{r}_1)d\vec{r}_1 - (1-2a)\int\int w_A(\vec{r}_1)w_A(\vec{r}_2) \\ &\qquad \times \left[\Gamma(\vec{r}_1,\vec{r}_2) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_2)\rho^s(\vec{r}_2;\vec{r}_1)\right] d\vec{r}_1 d\vec{r}_2 \\ &\quad -\frac{1}{2}\int\int w_A(\vec{r}_1)w_A(\vec{r}_2)\ \left[\Gamma(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_1)\rho^s(\vec{r}_2;\vec{r}_2)\right] d\vec{r}_1 d\vec{r}_2 \end{aligned}\tag{9.8}$$

and

$$\begin{aligned}\langle \hat{S}^2\rangle_{AB} = &-(1-2a)\int\int w_A(\vec{r}_1)w_B(\vec{r}_2)\ \left[\Gamma(\vec{r}_1,\vec{r}_2) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_2)\rho^s(\vec{r}_2;\vec{r}_1)\right] d\vec{r}_1 d\vec{r}_2 \\ -\frac{1}{2}\int\int w_A(\vec{r}_1)w_B(\vec{r}_2)\ &\left[\Gamma(\vec{r}_1,\vec{r}_2;\vec{r}_2,\vec{r}_1) - \tfrac{1}{2}\rho^s(\vec{r}_1;\vec{r}_1)\rho^s(\vec{r}_2;\vec{r}_2)\right] d\vec{r}_1 d\vec{r}_2, \end{aligned}\tag{9.9}$$

leading to a *continuum* of decompositions of $\langle \hat{S}^2\rangle$ fulfilling all requirements. We have introduced an additional criterion to select one formulation from the *continuum*. The criterion is based on the beaviour of eqs. (9.8) and (9.9) when applied to a one-electron system. The value $\langle \hat{S}^2\rangle = 3/4$ is an intrinsic property of the electron. Therefore, the density distribution of the $\langle \hat{S}^2\rangle$ for a single electron should simply be 3/4 times the electron density so that

$$\langle \hat{S}^2\rangle = \frac{3}{4}\int \rho(\vec{r})d\vec{r}. \tag{9.10}$$

For a single electron system (or ROHF systems with a single unpaired electron) eq. (9.7) reduces to eq. (9.10) **only** if $a = 3/4$. Furthermore, a=3/4 is the only value for which $\langle \hat{S}^2 \rangle_A$ is always positive for a simple FCI model of a minimal basis-set H_2 system. It is important to see that no diatomic terms will appear for a one-electron system, avoiding spurious self-coupling of the single electron. The additional criterion reads as (v) no two-center terms should appear in the case of single-electron systems.
Numerical results also indicate the $a = 3/4$ formula shows the smallest $\langle \hat{S}^2 \rangle$ values for pure covalent systems at equilibrium distances (see table (9.1)).

Table 9.1: CASSCF(2,4) atomic $\langle \hat{S}^2 \rangle_A$ values for the H_2 molecule at optimized geometries for several atomic definitions and values of the parameter a.

Atom in Molecule	a	cc-pVDZ	cc-pVTZ	cc-pVQZ
Becke	0	0.109	0.111	0.109
	3/8	0.072	0.073	0.073
	1/2	0.059	0.060	0.061
	3/4	**0.034**	**0.034**	**0.036**
Hirshfeld	0	0.122	0.124	0.122
	3/8	0.079	0.080	0.080
	1/2	0.064	0.065	0.066
	3/4	**0.035**	**0.036**	**0.038**
QTAIM	0	0.059	0.060	0.058
	3/8	0.044	0.044	0.045
	1/2	0.039	0.039	0.040
	3/4	**0.028**	**0.029**	**0.031**

Results in table (9.1) also reveal that the $\langle \hat{S}^2 \rangle$ values for $a = 3/4$ are almost basis set independent. Additionally, the effect of the AIM definition on the values $\langle \hat{S}^2 \rangle$ is very small for the preferred selection of parameter a. The local spin analysis has been applied to open shell systems (doublets and triplets) and to closed shell systems treated at the correlated level. Interestingly, the $\langle \hat{S}^2 \rangle$ values obtained with $a = 3/4$ are the smallest among the existing formulations when applied to genuine diamagnetic systems treated at the correlated level.

The decomposition of $\langle \hat{S}^2 \rangle$ can also be performed under the framework of the Hilbert-space analysis. Within this framework numerical integrations are no longer needed. Thus, the decomposition is in this sense "exact" and could be rather easily implemented in most quantum chemistry codes.

There is, however, an apparent ambiguity in decomposing two-electron quantities in the framework of Hilbert-space analysis that affects directly the decomposition of $\langle \hat{S}^2 \rangle$. To illustrate this ambiguity we have used the definition of the Mayer-Wiberg bond order for closed-shells

$$B_{AB} = \sum_{\mu \in A, \sigma \in B} (\mathbf{DS})_{\mu\sigma} (\mathbf{DS})_{\sigma\mu}. \tag{9.11}$$

Inspecting the expression 9.11 one can see that both of the subscripts (μ and σ), that serve for subdividing the quantity into diatomic contributions, come from the *"ket"* part of the overlap integral. But, in principle, any expression that contains products of two (or more) overlap matrices could also be subdivided in atomic and diatomic terms by taking one subscript from "bra" and another from "ket" (We refer to this convention as "interchanged" indices) . In the case of the bond order this leads to the definition of the bond order by Sato and Sakaki

$$B'_{AB} = \sum_{\mu \in A} \sum_{\nu \in B} (\mathbf{SDS})_{\mu\nu} D_{\nu\mu}. \tag{9.12}$$

Numerical results obtained by using eq. (9.12) are much less "chemical" than those obtained with the original Mayer-Wiberg expression (eq. (9.11)). By using the formalism of the effective atomic densities, one can see that our proposed decomposition of $\langle \hat{S}^2 \rangle$ within the Hilbert-space framework is naturally consistent with that of the original Mayer-Wiberg BO or Mulliken population analysis. The one- and two-center contributions to $\langle \hat{S}^2 \rangle$ can be written as

$$\begin{aligned} \langle \hat{S}^2 \rangle_A = & \frac{3}{4} \int u_A(\vec{r}_1) d\vec{r}_1 + \frac{1}{2} \iint \left[\Gamma_{AA}(\vec{r}_1, \vec{r}_2) - \tfrac{1}{2} \rho_A^s(\vec{r}_1; \vec{r}_2) \rho_A^s(\vec{r}_2; \vec{r}_1) \right] d\vec{r}_1 d\vec{r}_2 \\ & - \frac{1}{2} \int \int \left[\Gamma_{AA}(\vec{r}_1, \vec{r}_2; \vec{r}_2, \vec{r}_1) - \tfrac{1}{2} \rho_A^s(\vec{r}_1; \vec{r}_1) \rho_A^s(\vec{r}_2; \vec{r}_2) \right] d\vec{r}_1 d\vec{r}_2, \end{aligned} \tag{9.13}$$

and

$$\langle \hat{S}^2 \rangle_{AB} = \frac{1}{2} \iint \left[\Gamma_{AB}(\vec{r}_1, \vec{r}_2) - \tfrac{1}{2} \rho^s_A(\vec{r}_1; \vec{r}_2) \rho^s_B(\vec{r}_2; \vec{r}_1) \right] d\vec{r}_1 d\vec{r}_2$$
$$-\frac{1}{2} \int \int \left[\Gamma_{AB}(\vec{r}_1, \vec{r}_2; \vec{r}_2, \vec{r}_1) - \tfrac{1}{2} \rho^s_A(\vec{r}_1; \vec{r}_1) \rho^s_B(\vec{r}_2; \vec{r}_2) \right] d\vec{r}_1 d\vec{r}_2, \quad (9.14)$$

where the *effective* atomic spin density is written as

$$\rho^s_A(\vec{r}; \vec{r}') = \sum_{\mu \in A, \nu} P^s_{\mu\nu} \chi^*_\nu(\vec{r}) \chi_\mu(\vec{r}'), \quad (9.15)$$

the atomic density of effectively unpaired electrons is

$$u_A(\vec{r}) = \sum_{\mu \in A, \nu} \left(2D_{\mu\nu} - [DSD]_{\mu\nu} \right) \chi^*_\nu(\vec{r}) \chi_\mu(\vec{r}), \quad (9.16)$$

and the atomic/diatomic spin-less cumulant reads as

$$\Gamma_{AB}(\vec{r}_1, \vec{r}_2; \vec{r}_1{}', \vec{r}_2{}') = \sum_{\mu \in A, \sigma \in B, \nu, \lambda} \Gamma_{\mu\sigma\nu\lambda} \, \chi^*_\nu(\vec{r}_1) \chi^*_\lambda(\vec{r}_2) \chi_\mu(\vec{r}_1{}') \chi_\sigma(\vec{r}_2{}'). \quad (9.17)$$

This leads to the final expressions for the atomic terms

$$\begin{aligned} \langle \hat{S}^2 \rangle_A &= \frac{3}{4} \sum_{\mu \in A} \left(2 [DS]_{\mu\mu} - [DSDS]_{\mu\mu} \right) \\ &\quad -\frac{1}{4} \sum_{\mu,\nu \in A} [P^s S]_{\mu\nu} [P^s S]_{\nu\mu} + \frac{1}{4} \sum_{\mu,\nu \in A} [P^s S]_{\mu\mu} [P^s S]_{\nu\nu} \\ &\quad +\frac{1}{2} \sum_{\mu,\sigma \in A} \sum_{\nu,\lambda} \left(\Gamma_{\mu\sigma\nu\lambda} - \Gamma_{\mu\sigma\lambda\nu} \right) S_{\lambda\sigma} S_{\nu\mu}, \end{aligned} \quad (9.18)$$

and similarly for the diatomic spin components

$$\begin{aligned} \langle \hat{S}^2 \rangle_{AB} &= -\frac{1}{4} \sum_{\mu \in A, \nu \in B} [P^s S]_{\mu\nu} [P^s S]_{\nu\mu} + \frac{1}{4} \sum_{\mu \in A, \nu \in B} [P^s S]_{\mu\mu} [P^s S]_{\nu\nu} \\ &\quad +\frac{1}{2} \sum_{\mu \in A, \sigma \in B} \sum_{\nu,\lambda} \left(\Gamma_{\mu\sigma\nu\lambda} - \Gamma_{\mu\sigma\lambda\nu} \right) S_{\lambda\sigma} S_{\nu\mu}. \end{aligned} \quad (9.19)$$

Table 9.2: Atomic local spin values calculated at the CASSCF(2,4) level for the H_2 molecule at interatomic distance $R_{H\text{-}H}$=0.746 Å for several basis sets.

Basis set	$\langle \hat{S}^2 \rangle_H$	$\langle \hat{S}^2 \rangle_H^{int}$	$\langle \hat{S}^2 \rangle_H^{L}$	$\langle \hat{S}^2 \rangle_H^{3D}$
cc-pVDZ	0.036	0.060	0.028	0.034
cc-pVTZ	0.035	0.051	0.025	0.034
cc-pVQZ	0.039	0.066	0.027	0.036
cc-pVTZ (6d)[a]	0.035	0.052	0.027	0.034
cc-pVQZ (6d 10f)[b]	0.043	0.089	0.028	0.036
aug-cc-pVDZ	0.038	0.069	0.025	0.034
aug-cc-pVTZ	0.020	0.031	0.026	0.036
aug-cc-pVQZ	0.035	0.044	0.026	0.036
d-aug-cc-pVQZ	0.045	0.130	0.024	0.034
aug-cc-pVTZ (6d)	0.046	0.115	0.027	0.036
aug-cc-pVQZ (6d 10f)	0.055	0.192	0.026	0.036
d-aug-cc-pVQZ (6d 10f)	0.095	0.891	0.023	0.034

[a] Cartesian d functions used.

[b] Cartesian d and f functions used.

Moreover, it is important to note that a mapping between the decomposition of $\langle \hat{S}^2 \rangle$ in the Hilbert- and 3D-space is only possible using the original Mayer-Wiberg convention.

To assess the numerical effects on the use of different summation schemes in Hilbert-space analysis we have studied the H_2 molecule (see table (9.2)). We have also included values for Löwdin ($\langle \hat{S}^2 \rangle_H^L$) and 3D-space analysis ($\langle \hat{S}^2 \rangle_H^{3D}$) for comparison. $\langle \hat{S}^2 \rangle_H$ and $\langle \hat{S}^2 \rangle_H^{int}$ correspond to the conventional and "interchanged indices" Hilbert-space results, respectively. In the Löwdin basis the overlap matrix is diagonal so the schemes with conventional and "interchanged" indices are the same.

The local spin values using eq. (9.18), that is, with the conventional Hilbert-space scheme, are close to zero as expected for diamagnetic molecules. No basis set dependencies have been detected in this case and the results are very similar to those obtained in the framework of the 3D-space analysis. Using the alternative "interchanged indices" formula the results exhibit a stronger basis set dependence. The $\langle \hat{S}^2 \rangle_H^L$ and $\langle \hat{S}^2 \rangle_H^{3D}$ values are almost basis set independent.

The preferred decomposition, eqs. (9.18) and (9.19), has been applied to a

series of singlet molecules for illustration. For these systems the overall $\langle \hat{S}^2 \rangle$ is 0 but small atomic local spin contributions appear induced by electron correlation. Only in the case of an antiferromagnetic coupling large atomic local spins should appear.

Table 9.3: Atomic $\langle \hat{S}^2 \rangle_{\mathrm{C}}$ components for acetylene molecule computed at the CISD level of theory with different basis sets.

Basis set	$\langle \hat{S}^2 \rangle_{\mathrm{C}}$	$\langle \hat{S}^2 \rangle_{\mathrm{C}}^{L}$	$\langle \hat{S}^2 \rangle_{\mathrm{C}}^{3D}$
STO-3G	0.012	0.179	0.159
6-31G	-0.074	0.119	0.114
6-31G**	-0.136	0.082	0.084
6-31G**(6d)	-0.139	0.080	0.083
6-311G	0.030	0.119	0.113
6-311G**	-0.118	0.085	0.085
6-311G**(6d 10f)	-0.143	0.091	0.084
cc-pVDZ	-0.162	0.078	0.085
cc-pVTZ	-0.271	0.087	0.079
cc-pVTZ(6d 10f)	-0.146	0.104	0.078
aug-cc-pVDZ	0.729	0.078	0.087
aug-cc-pVTZ	-2.926	0.092	0.078

[a] Cartesian d functions used.
[b] Cartesian d and f functions used.

The most striking result of our analysis is a negative local value of $\langle \hat{S}^2 \rangle_{\mathrm{C}}$ for the acetylene molecule. The truncated CISD wave function is not responsible for this odd result because at the same level of theory but within the 3D-space framework $\langle \hat{S}^2 \rangle_{\mathrm{C}} = 0.083$. We have checked the basis set effects on this system; the results are gathered in table (9.3). The $\langle \hat{S}^2 \rangle_{\mathrm{C}}$ values are negative for all the basis sets except for STO-3G, 6-311G and aug-cc-pVDZ. A completely meaningless value of -2.926 appears for the aug-cc-pVTZ basis set. Negative values are obtained for both small and relatively large basis set, with and without polarization functions, so it is difficult to find a general explanation for this odd behavior. The $\langle \hat{S}^2 \rangle_{\mathrm{C}}^{L}$ and $\langle \hat{S}^2 \rangle_{\mathrm{C}}^{3D}$ values are always small, positive and show a relative very small dependence. Similar odd behaviour has been observed for the values of delocalization indices. There seems to be a

shortcoming when combining Hilbert-space analysis and genuine two electron quantities (*i. e.* using the 2-RDM) that deserves further analysis.

9.3 The local spin analysis: relationship with chemical bonding and radical character

Once a unique formulation of the local spin has been established for any AIM definition, we are in a position to gain insight on the interpretation of the $\langle \hat{S}^2 \rangle_A$ and $\langle \hat{S}^2 \rangle_{AB}$ values and their relation with the concept of a covalent bond. For that purpose we have studied two ideal systems. The first one is a two-electron homonuclear diatomic system in the singlet state treated at the full configuration interaction (FCI) minimal basis level, consisting of two hydrogen-like atoms labeled A and B. For such a system, the elements of the atomic natural spinorbital overlap matrix within a general 3D-space framework are

$$\begin{aligned} S_{ii}^A &= S_{ii}^B = 1/2 \\ S_{ij}^A &= -S_{ij}^B = \delta. \end{aligned} \tag{9.20}$$

In the dissociation limit $\delta = 1/2$ but at interatomic distances this value can slightly decrease. The analytical expression of the bond order (BO), the atomic number of effectively unpaired electrons u_A, $\langle \hat{S}^2 \rangle_A$, and $\langle \hat{S}^2 \rangle_{AB}$ can be written in terms of δ and the occupation of the antibonding natural spinorbital n as

$$BO = 1 - 2n(1-n)(1+4\delta^2) \tag{9.21}$$

$$u_A = 4n(1-n) \tag{9.22}$$

$$\langle \hat{S}^2 \rangle_A = -\langle \hat{S}^2 \rangle_{AB} = \frac{3}{2} n(1-n)(1+4\delta^2). \tag{9.23}$$

Let us consider now the same two-electron system but in the triplet state. In this case, the BO, the atomic number of effectively unpaired electrons u_A, $\langle \hat{S}^2 \rangle_A$, and $\langle \hat{S}^2 \rangle_{AB}$ depends only upon δ, namely

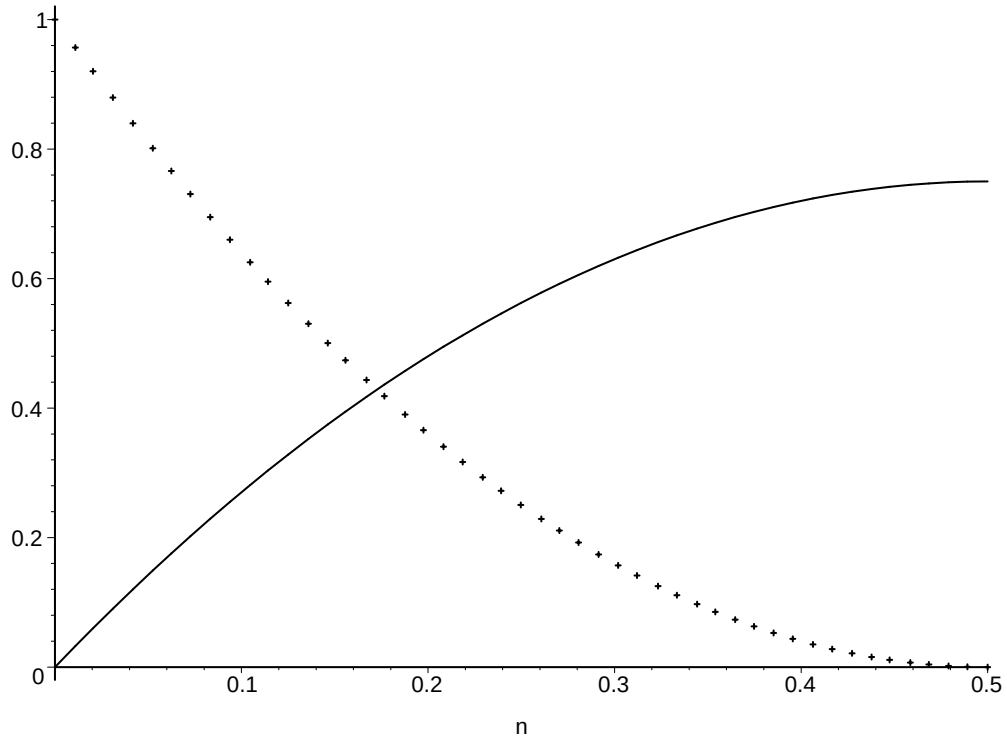

Figure 9.5: Evolution of the bond order (dotted) and $\langle \hat{S}^2 \rangle_A$ (solid) for the two-electron singlet model system (assuming $\delta = 1/2$) with respect to the natural spin orbital occupation, n.

$$BO = 1 - 4\delta^2 \tag{9.24}$$

$$\langle \hat{S}^2 \rangle_A = \frac{7}{8} - \frac{1}{2}\delta^2 \tag{9.25}$$

$$\langle \hat{S}^2 \rangle_{AB} = \frac{1}{8} + \frac{1}{2}\delta^2. \tag{9.26}$$

The information provided by these two ideal systems can be summarized as follows (see also fig. (9.5)). For the singlet state case described with a single-determinant wave function ($n = 0$) the local spin analysis yields $\langle \hat{S}^2 \rangle_A = \langle \hat{S}^2 \rangle_{AB} = 0$. By including electron correlation (via a multireference wave function or using the unrestricted broken symmetry (BS) approach) $\langle \hat{S}^2 \rangle_A$ increases monotonically up to a maximum value of 3/4, the $\langle \hat{S}^2 \rangle$ value of an isolated electron. In fact, one can easily see that for this model system $\langle \hat{S}^2 \rangle_A = \frac{3}{4}(1 - BO)$, that is, the extent of local spin is proportional to the deviation from the perfect covalent bond picture. The diatomic contribution $\langle \hat{S}^2 \rangle_{AB} < 0$ reflects the coupling of the two spins, but its limiting value depends upon the nature of the wave function. A proper multireference wave function renders the entanglement of the spins, giving $\langle \hat{S}^2 \rangle_{AB} = -3/4$ in the limit of two perfectly localized spins, whereas for a BS wave function one can see that $\langle \hat{S}^2 \rangle_{AB} = -1/4$.

For the system in triplet state the local spin analysis gives non-zero $\langle \hat{S}^2 \rangle_A$ and $\langle \hat{S}^2 \rangle_{AB}$ contributions, as expected for a system with an overall non-zero spin. In the limit of two perfectly localized parallel spins on centers A and B, $\langle \hat{S}^2 \rangle_A = 3/4$ and $\langle \hat{S}^2 \rangle_{AB} = 1/4$. When the two spins are not perfectly localized, the value of $\langle \hat{S}^2 \rangle_A$ increases, indicating a given probability of finding the two parallel electrons on each atomic center.

In order to put into perspective the results of the local spin analysis of real molecules at their ground state geometries, it is worth analyzing how the components of the local spin behave in dissociation. To simplify the analysis we have written eq. (9.13) and (9.14) in the compact form

$$\langle \hat{S}^2 \rangle_A = \frac{3}{4} u_A + \Lambda_{AA} + \Lambda'_{AA} \tag{9.27}$$

and

$$\langle \hat{S}^2 \rangle_{AB} = \Lambda_{AB} + \Lambda'_{AB}, \tag{9.28}$$

where the contributions can be most generally written in terms of the *effective* atomic densities as

$$\begin{aligned}
u_A &= \frac{3}{4} \int u_A(\vec{r}_1) d\vec{r}_1 \\
\Lambda_{AA} &= \frac{1}{2} \iint \left[\Gamma_{AA}(\vec{r}_1, \vec{r}_2) - \tfrac{1}{2} \rho^s_A(\vec{r}_1; \vec{r}_2) \rho^s_A(\vec{r}_2; \vec{r}_1) \right] d\vec{r}_1 d\vec{r}_2 \\
\Lambda'_{AA} &= -\frac{1}{2} \int \int \left[\Gamma_{AA}(\vec{r}_1, \vec{r}_2; \vec{r}_2, \vec{r}_1) - \tfrac{1}{2} \rho^s_A(\vec{r}_1; \vec{r}_1) \rho^s_A(\vec{r}_2; \vec{r}_2) \right] d\vec{r}_1 d\vec{r}_2. \\
\Lambda_{AB} &= \frac{1}{2} \iint \left[\Gamma_{AB}(\vec{r}_1, \vec{r}_2) - \tfrac{1}{2} \rho^s_A(\vec{r}_1; \vec{r}_2) \rho^s_B(\vec{r}_2; \vec{r}_1) \right] d\vec{r}_1 d\vec{r}_2 \\
\Lambda'_{AB} &= -\frac{1}{2} \int \int \left[\Gamma_{AB}(\vec{r}_1, \vec{r}_2; \vec{r}_2, \vec{r}_1) - \tfrac{1}{2} \rho^s_A(\vec{r}_1; \vec{r}_1) \rho^s_B(\vec{r}_2; \vec{r}_2) \right] d\vec{r}_1 d\vec{r}_2.
\end{aligned} \tag{9.29}$$

For a singlet diatomic molecule dissociating into two doublet moieties a proper wave function would yield $u_A = 1$. Therefore, the contribution of u_A fully accounts for the local spin of the atoms in dissociation. This is not the general trend for diatomic molecules at dissociation. In fact, the expected contribution from $\Lambda_{AA} + \Lambda'_{AA}$ in dissociation is in general $u_a(u_a - 1)/4$, vanishing for systems dissociating into singlet or doublet atoms. As in the case of the two ideal systems above, it is the $\langle \hat{S}^2 \rangle_{AB}$ term that accounts for the coupling between any existing localized spins.

We have performed the local spin analysis of a series of diatomic molecules in their ground state equilibrium distances, including the challenging and polemic[93–95] C_2 molecule. In order to gain deeper insight into the nature of the local atomic spin and the diatomic spin components, these values have been further decomposed according to the spatial symmetry of the MOs. In the case of the one electron quantities like the number of effectively unpaired electrons such decomposition leads to additive σ and π contributions. For genuine two-electron quantities like the cumulants – and hence both $\langle \hat{S}^2 \rangle_A$ and $\langle \hat{S}^2 \rangle_{AB}$ terms– one has σ-σ,π-π and also σ-π contributions.

In C_2, the number of effectively unpaired electrons at each C atom is 1.10.

It arises from two similar σ (0.66) and π (0.43) contributions. We have seen that a large value of the number of unpaired electrons is a necessary but not sufficient condition to picture the system as a diradical (*e.g.* the case of Be_2). However, the $\langle \hat{S}^2 \rangle_A$ value is also very large in this case (0.81), the highest among all singlet states considered, and even slightly larger than the expected number for a perfectly localized spin (0.75). It is also the only case where the σ-π cross-term contributes to increase the local spin. The effect of the cross-terms is most notorious in the diatomic spin component, where all σ-σ (-0.35), π-π (-0.20) and σ-π (-0.26) terms contribute significantly to the overall -0.81 value. We have also considered other CASSCF wave functions where only σ or π orbitals enter in a (4,4) active space. The $\langle \hat{S}^2 \rangle_{AB}$ values obtained, (-0.38 for an all σ and -0.11 for the all π active spaces) are not too far from the $\langle \hat{S}^2 \rangle_{AB}^{\sigma-\sigma}$ and $\langle \hat{S}^2 \rangle_{AB}^{\pi-\pi}$ contributions mentioned above, certainly not large enough to clearly label the system neither as a σ nor π diradical. However, the two shares put together and enhanced by a σ-π interplay result in a genuine diradicaloid system. The fact that this can be only observed upon inclusion of electron correlation (*i.e.,* when the antibonding π orbitals also come into play) may suggest a sort of superexchange mechanism, where the π orbitals mediate the exchange interaction between two localized σ hybrid orbitals sitting on each C atom. The values of the local spin analysis for the heteroatomic system CN^+ are strikingly similar to those of their isoelectronic molecule C_2. For CB^- and BN the local spins slightly decrease with respect to the dicarbon values. The heavier homologue of C_2, that is, the Si_2 molecule, presents a local spin values closer to O_2 than to either C_2 or BN.

On the other hand, the study also reveals that the presence of local spin flags deviation from the classical bonding prototypes. The next step is to study the applicability of the $\langle \hat{S}^2 \rangle$ decomposition analysis to characterize radical centers in organic molecules. For this purpose the local spin analysis is tested against several indexes for diradical character for the simple dissociating curve of H_2.

The spin analysis and the number of effectively unpaired electrons capture and quantify diradical character in a similar manner for the model. Even though these two descriptors are the only ones that can be easily generalized to polyradicals, the number of effectively unpaired electrons exhibits some

undesirable features. First, it has an upper bound of $2N$, where N is the total number of electrons. Thus the number of effectively unpaired electrons may be larger than the actual number of electrons in the system. The second drawback, noted by Staverov and Davidson some years ago, is related to the puzzling result found in the dissociation of O_2 in its $^3\Sigma_g^-$ ground state into two triplet ^3P oxygen atoms. A value of 5 effective unpaired electrons was found at the dissociation limit instead of the expected value of 4, i.e. 2 on each triplet O atom. We have further explored this paradigmatic system by considering for a number of different electronic states the dissociation of O_2 into two O atoms. We have seen that the local spin values *always* yield the expected values for the dissociating oxygen atoms, *i.e.*, $\langle \hat{S}^2 \rangle_{\rm O} = 2$ for ^3P and $\langle \hat{S}^2 \rangle_{\rm O} = 0$ for ^1D atomic oxygen states, respectively. The diatomic spin components also differentiate when the two oxygen triplets are coupled as a singlet, like in the $^1\Delta_g$ state for which $\langle \hat{S}^2 \rangle_{\rm O,O} = -2$, or as a triplet, like in the $^3\Sigma_g^-$ state, for which $\langle \hat{S}^2 \rangle_{\rm O,O} = -1$ is obtained. Thus, the local spin analysis appears to be a more suitable tool than the number of effectively unpaired electrons when it comes to the formal breaking of more than one bond. This is in essence because the $\langle \hat{S}^2 \rangle_A$ terms include contributions from the cumulant of the second order-density matrix, whereas the number of effectively unpaired electrons is obtained only from the first-order density matrix. Accordingly, the quantification of polyradical character will be better accomplished by making use of the descriptors obtained from the local spin analysis.

We have introduced a general measure of k-radical character making use of the $\langle \hat{S}^2 \rangle_A$ values obtained for an actual molecular system as

$$\Delta^{(k)} = \sqrt{\frac{\sum_A (\langle \hat{S}^2 \rangle_A - \langle \hat{S}^2 \rangle_A^{id})^2}{n}}, \tag{9.30}$$

where $\langle \hat{S}^2 \rangle_A^{id}$ represents the atomic ideal values and n is the total number of atoms/fragments considered in the local spin analysis. The smaller the $\Delta^{(k)}$ value the closer the system is to a reference picture of k perfectly localized spin centers. Note that in order to be $\Delta^{(k)}$ an index of most general applicability it only utilizes the atomic contributions of the local spin analysis. Indeed, we have already seen that the *ideal* value for the diatomic terms depends upon

Table 9.4: k-radical character index, $\Delta^{(k)}$, for *ortho*-, *meta*-, and *para*-benzine.

	Singlet		Triplet
Molecule	$\Delta^{(0)}$	$\Delta^{(2)}$	$\Delta^{(2)}$
o-benzyne	0.18	0.28	0.14
m-benzyne	0.28	0.19	0.13
p-benzyne	0.49	0.12	0.12

the particular electronic state (singlet, triplet, ...) and also the type of wave function. This indicator can be computed for any system in any state, and its value depends upon the particular choice of the $\langle \hat{S}^2 \rangle_A^{id}$ terms.

For diradical or diradicaloid species in singlet-state one can calculate both $\Delta^{(0)}$ and $\Delta^{(2)}$ values. The former measures the average deviation from a nonradical closed-shell picture, and the latter indicates the deviation from a perfect diradical. The trends of the $\Delta^{(0)}$ and $\Delta^{(2)}$ values along the series of singlet ortho-, meta- and para-benzyne are very illustrative (see table (9.4)). For o-benzyne $\Delta^{(0)} = 0.18$ and $\Delta^{(2)} = 0.28$, indicating a smaller deviation of this species with respect to a closed-shell picture. For m-benzyne the situation is just the opposite, and the system is better identified as a diradical. For p-benzyne the $\Delta^{(2)}$ value is very small (0.12) and much smaller than the $\Delta^{(0)}$ one (0.49), which is consistent with a diradical picture. Thus, both $\Delta^{(0)}$ and $\Delta^{(2)}$ values are able to reproduce the assumed trend ortho $<$ meta $<$ para of the diradical character of benzyne isomers.[96,97] For triplet states only the $\Delta^{(2)}$ values are meaningful. The values are quite small and do not differ too much from one isomer to another. The trend along the series is the same as for the singlet states *i.e.,* triplet para-benzyne is more diradical than ortho-benzyne. Several other diradicaloid systems have also been studied. We have found that diphospadiboretane and its diaza-analogue are best described as closed-shell and delocalized-spin species in their singlet and triplet states, respectively. The local spin analysis performed on [2,2,2]propellane also confirm its nonradical nature, even in the absence of the central C-C bond.

We have illustrated the capability of the local spin analysis to characterize triradicals analyzing the simplest model of a triradical *i.e.,* three H atoms de-

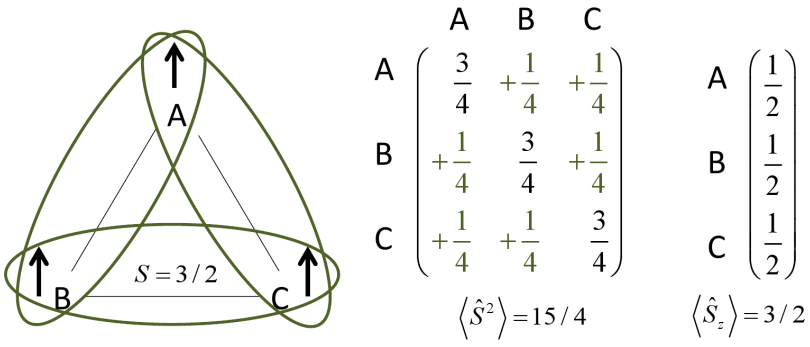

Figure 9.6: Pictorial representation of the local spin analysis for perfectly localized spins in quartet states.

scribed with a minimal basis placed at the vertices of a triangle. For such a system one can have a quartet state with $S = 3/2$ and two doublet states with $S = 1/2$. The perfectly localized spins situation is achieved for large H-H distances. A pictorial representation of the results of the local spin analysis obtained at the dissociation limit for the quartet state is given in fig. (9.6). The $\langle \hat{S}^2 \rangle_A$ and $\langle \hat{S}^2 \rangle_{AB}$ values and the atomic $< \hat{S}_z >$ values (half of the atom-condensed spin-density) are also shown. The local spin, $\langle \hat{S}^2 \rangle_A$, on each hydrogen atom is 3/4, consistent with that of a single electron. The diatomic spin contributions $\langle \hat{S}^2 \rangle_{AB}$ are equal to +1/4 for any A≠B, indicating the parallel arrangement of the local spins. The spin density on each H atom equals to 1, and so is the number of effectively unpaired electrons. When the spins are not perfectly localized, the values of $\langle \hat{S}^2 \rangle_A$ increase, indicating partial high-spin contributions on the centers (triplet and quartet, in this case) due to the non-negligible probability of finding more than one electron on a given atom. Consequently, the $\langle \hat{S}^2 \rangle_{AB}$ terms decrease.

In the case of the doublet states one can encounter two limiting situations for perfectly localized spins, represented in fig. (9.7). The upper panel describes a situation in which two of the perfectly localized spins are coupled as a singlet (centers B and C), whereas the third spin does not show any coupling with them, *i.e.*, $\langle \hat{S}^2 \rangle_{AB} = \langle \hat{S}^2 \rangle_{AC} = 0$. The spin density is zero for the centers B

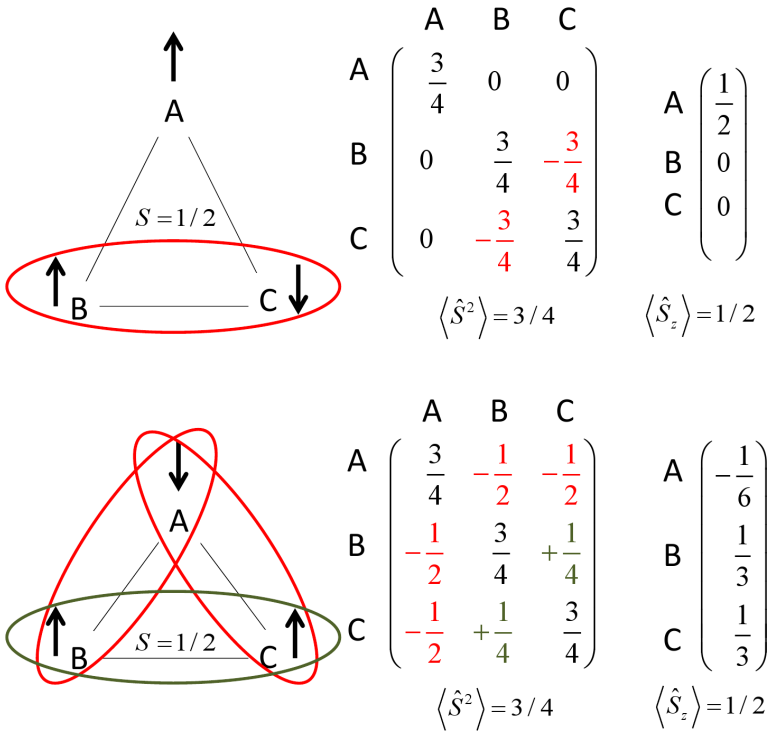

Figure 9.7: Pictorial representation of the local spin analysis for perfectly localized spins in doublet states.

and C and 1 for the center A. This represents a case in which the triradical is best pictured as a diradical plus an isolated radical center.

Another situation is depicted in the lower panel of fig. (9.7). In this case, the three perfectly localized spins interact with each other. The spins on centers B and C exhibit parallel arrangement, as indicated by the $\langle \hat{S}^2 \rangle_{BC} = +1/4$ value, and are antiparallel with the local spin on center A. The diatomic spin contributions $\langle \hat{S}^2 \rangle_{AB} = \langle \hat{S}^2 \rangle_{AC}$ are equal to -1/2. The negative sign is an indication that the spins on the centers are antiparallel but the actual value differs from the -3/4 that one obtains for a pair of isolated spins coupled as a singlet. The atomic spin densities are non-zero, but can not be trivially associated with three perfectly localized spins: the spin density on center A amounts to -1/3 whereas that of centers B and C is 2/3, for an overall value of 1. This illustrates once again the limitation of use of the spin density for the description of local spins. The simple pictorial representation of fig. (9.7) can only be derived from the local spin analysis.

Thus, the magnitude and sign of the diatomic spin terms differentiate between doublet and quartet states. Moreover, they also reveal that two patterns of spins distribution may occur in the lowest spin state. One type is best characterized as a system in which one can distinguish a singlet diradical and an additional isolated radical center. In the other one the three spin centers exhibit interactions with each other.

The spin properties of a number of all σ, all π and σ–π triradicals namely trimetylenebenzene (TMB),[98–103] tridehydrobenzene isomers (TDB),[103–106] and dehydro-m-xylylene (DMX)[103,107,108] have been characterized in detail in their lowest-lying doublet and quartet electronic states. Among all triradical systems in their lowest-lying doublet states, the $^2B_1$ state of TMB, 1,2,4-TDB, and 1,3,5-TDB have been characterized a singlet diradical + an isolated radical center. The others doublet state triradicals have been characterized as triradicals with three spin centers interacting with each other.

9.4 Effective atomic orbitals: developments and applications

The most appropriate entities to characterize the state of the atom within the molecule are the effective atomic orbitals. In this approach one can retrieve the electronic configuration of atoms in molecules from wave function analysis. We have implemented for the first time the eff-AOs in the framework of Bader's topological atoms. The main visual feature of the hybrids is that they are completely confined within the atomic domain, i.e., in this case they are strictly cut at the boundary limit of Bader's topological atom. A 2s-type eff-AO of the hydrogen atom in methane is depicted in fig. (9.8), for illustration.

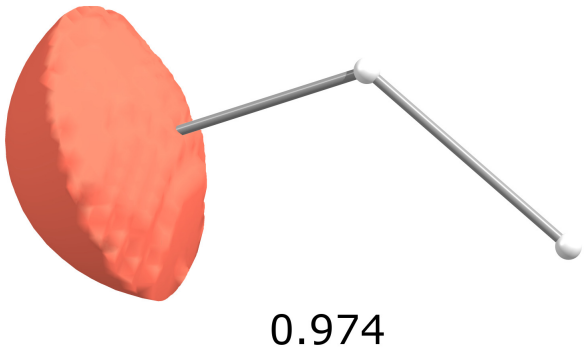

Figure 9.8: 2s-type eff-AO for H atom in methane and its occupation number.

Numerical results on the HCNO molecule and the series $CH_3SO_xCH_3$, x=0,1,2 confirm that the number of eff-AOs with a significant occupation number on each atom is limited and usually coincide with the number of orbitals in a minimal basis set. However, hypervalent atoms, such as nitrogen and sulfur in the cases studied, may exhibit slightly occupied eff-AOs beyond the valence shell.

On the other hand, we have shown that the original MOs of a calculation can be written as a linear combination of the eff-AOs with non-zero eigenvalues of different atoms as

Table 9.5: Accuracy of the SVD procedure for an Ala molecule with two basis sets (number of basis functions in parentheses). N_e indicates the number of effective atomic orbitals with occupation number above the threshold. $\bar{\delta}$ and δ_{max} are the average and maximum error in the atomic population values after the SVD procedure.

	6-31++G(d,p) (156)	cc-pVTZ (278)
	occ. number > 0.001	
N_e	114	119
$\bar{\delta}$	$8.9 \cdot 10^{-4}$	$8.8 \cdot 10^{-4}$
δ_{max}	$1.9 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$
	occ. number > 0.01	
N_e	69	71
$\bar{\delta}$	$7.6 \cdot 10^{-3}$	$7.3 \cdot 10^{-3}$
δ_{max}	$1.7 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$
	occ. number > 0.1	
N_e	37	37
$\bar{\delta}$	$3.2 \cdot 10^{-2}$	$3.1 \cdot 10^{-2}$
δ_{max}	$9.8 \cdot 10^{-2}$	$8.9 \cdot 10^{-2}$

$$\varphi_i(\vec{r}) = \sum_{A}^{N_{at}} \sum_{\mu=1}^{n_A} U_{i\mu}^{A*} \sqrt{\lambda_\mu^A} \chi_\mu^A(\vec{r}) \quad . \tag{9.31}$$

The atomic populations and the classical Wiberg BO calculated by the Hilbert-space analysis in the basis of the eff-AOs are equal to those computed by the QTAIM analysis of the total density. Moreover, taking advantage of the properties of the hybrids, the MOs can be expanded, to a good approximation, as linear combination of a reduced number (only those with significant occupation numbers) of eff-AOs. The new LCAO coefficients on the eff-AO basis were obtained by a singular value decomposition (SVD) of a rectangular matrix containing the overlap elements of the MOs with the reduced set of eff-AOs. Remarkably, highly accurate population analysis can be produced by the approximate MOs expanded in the reduced set of eff-AOs (see table (9.5)).

From the analysis of the effective atomic orbitals one can derive atomic electron configurations for the atom in the molecule, by proper comparison of the occupation numbers of the eff-AOs. These electron configuration can be further used to infer effective oxidation states (EOS). A spin-resolved version of the eff-AO method is used instead, in order to conserve the number of alpha and beta electrons separately. The procedure is as simple as follows:

(i) Collect the alpha eff-AOs that are significantly populated for *all* centers.

(ii) Sort them according to decreasing occupation number.

(iii) Round the occupation numbers to 1 starting from the highest occupied eff-AO until the number of alpha electrons is reached, and do analogously for the beta part.

Thus, each (integer) alpha and beta electron is assigned to the atom where the given eff-AO belongs to, whereas the occupation of the remaining eff-AOs is rounded off to zero. By this process an (integer) electronic configuration is obtained for each atom. The effective oxidation state (EOS) of each atom is simply given by the difference between its atomic number and the number of alpha and beta electrons that have been assigned to it. Note that with this scheme the EOS of a given atom does not only depend upon the population of its eff-AOs but also on that of the remaining atoms, and of course on the total number of electrons of the system. Moreover, the occupation numbers of the last occupied, λ^{σ}_{LO}, and first unoccupied, λ^{σ}_{FU}, eff-AOs (before rounding to 1 and 0, respectively) indicate how well-suited the EOS are for the actual electronic distribution of the system. Ideally, they should be close to 1 and 0, respectively, but such values are only expected for non-interacting atoms. In practice the atoms share the electrons and this is reflected in their relative occupation numbers. Since EOS are determined by integer electrons, when λ^{σ}_{LO} and λ^{σ}_{FU} differ by more than *half* electron (*i.e.,* rounding up to one electron) the assignation of EOS is considered as undisputable. A "confidence" index for the EOS arising from the analysis can be simply defined as

$$C_{\sigma}(\%) = 100\ min(1, \lambda^{\sigma}_{LO} - \lambda^{\sigma}_{FU} + 1/2). \tag{9.32}$$

The alpha and beta values can be simply averaged to get an overall C(%) value.

Practical applications of the method may involve reaction intermediates or transition state structures of complexes exhibiting bulky ligands. When the number of atoms of the system is large, accidental degeneracies of the occupation numbers of the eff-AOs are likely to occur, which difficult the assignation of electrons. A better strategy is a hierarchical approach, by which molecular fragments are defined *before* the eff-AO analysis. That is, instead of eff-AOs we obtain effective fragment orbitals by using in eq. (1.70) fragment weight functions of the form

$$w_K(\vec{r}) = \sum_{i \in K} w_i(\vec{r}), \tag{9.33}$$

where the sum runs over all atoms of molecular fragment K. In TM complexes the fragments are identified with the metal atom, the ligands and the molecular species that may be present, such as reactants or explicit solvent molecules, if any. Then, the steps (i)-(iii) described above lead to a proper distribution of the electrons of the system among the different fragments. If necessary, the EOS of the individual atoms forming a fragment can be derived in a second step by computing their eff-AOs and distributing the alpha and beta electrons that have been assigned to the fragment.

The method can be applied in principle beyond transition metal complexes. As could be anticipated, numerical results show that the worst scenario for the method are species with highly apolar bonds such C-H and C-C. The C(%) values for a set of hydrocarbons are slightly above 50% basically because the population of the 1s-type eff-AO on H atoms is comparable to that of the carbon's valence hybrids. There is no clear cut distinction between proton and hydride assignation. These compounds indicate the limit of applicability of the method. Nevertheless, for hydrocarbons and such highly apolar compounds the OS are genuinely formal and have probably little relevance in most chemical applications. Of course an alternative that can be applied in these cases is the fragment approach discussed above. Methyl groups or $(CH)_n$ moieties in general may be chosen as fragments so that the apolar C-H bonds are not *cut* in the eff-AO analysis.

In order to check how dependent the EOS are on the particular atom in molecule definition used we have performed a systematic study for 10 octahedral Fe^{II} and Fe^{III} complexes. We have used different atomic definitions, namely Hilbert-space Mulliken's, Bader's QTAIM, and several fuzzy atom schemes such as Hirshfeld, Hirshfeld-Iterative and the simplest Becke atoms. The EOS obtained are independent of the particular atomic definition. The only exception is $FeCN_6^{3-}$ using Becke atoms, where the alpha population of a d-type orbital on the atom competes with that of the valence's hybrid on all N atoms (see table (9.6)). As a result, the EOS assignation reported as a C(%) value slightly below 50% in the alpha part. On the other hand, the values of the confidence index do differ significantly from one method to another. Best performing approaches (in terms of high C(%) values) are QTAIM, TFVC and Hirshfeld-Iterative. It is worth to remark that in the case of the Hilbert-space type approach (Mulliken) the occupation numbers of the eff-AOs are not strictly restricted to the $0 \leq \lambda_i^A \leq 2$ range so that 3D-space analysis appears to be more appropriate from a conceptual perspective.[37] Moreover, it is well known that these methods suffer from basis set dependencies. In our case we have used a medium-sized basis set (6-31G*) with marked atomic character. Mulliken-type results are comparable to Hirshfeld, but experience tells us that the Mulliken-type results reported here can not be extrapolated to a larger basis such as cc-pVTZ, and therefore are not recommended. Thus, we can conclude that for EOS assignation atomic definitions that better take into account bond polarization are the most appropriate. For its simplicity we recommend the use of the TVFC scheme.

The fragment approach has been tested for a number of isolated hexacoordinated [Fe(Pytacn)] complexes [a] that are involved in C-H catalytic hydroxylation cycles.[109] Pytacn is a tetradentate ligand and the other two positions can be occupied by oxygen-containing ligands such as aquo, hydroxo or oxo, formally considered as $H_2O^{(0)}$, $OH^{(1-)}$ and $O^{(2-)}$ species. The fragment approach yields in all cases the chemically expected oxidation states for the central Fe atom and the ligands. The main observation is that the C(%) values decrease as the OS of the central metal atom increases. The formal picture of the

[a] Pytacn stands for 1-(2-pyridylmethyl)-4,7-dimethyl-1,4,7-triazacyclononane.

Table 9.6: EOS, last occupied and first unoccupied eff-AOs and C(%) values for the $Fe(CN)_6^{3-}$ complex.

Atomic definition	Atom	EOS	$\lambda^{\alpha}_{LO}/\lambda^{\alpha}_{FU}$	$\lambda^{\beta}_{LO}/\lambda^{\beta}_{FU}$	$C_{\alpha}(\%)/C_{\beta}(\%)$
	Fe	3	0.896 / 0.327	0.871 / 0.283	
Mulliken	C	2	0.765 / 0.396	0.73 8 / 0.280	72 / 79
	N	-3	0.617 / -	0.569 / -	
	Fe	3	0.905 / 0.432	0.886 / 0.393	
Becke	C	2	0.596 / 0.350	0.607 / 0.364	47 / 51
	N	-3	0.406 / -	0.407 / -	
	Fe	3	0.832 / 0.358	0.812 / 0.319	
Hirshfeld	C	2	0.584 / 0.288	0.596 / 0.299	56 / 60
	N	-3	0.418 / -	0.419 / -	
	Fe	3	0.823 / 0.349	0.802 / 0.310	
Hirshfeld-Iterative	C	2	0.577 / 0.272	0.590 / 0.283	58 / 62
	N	-3	0.425 / -	0.426 / -	
	Fe	3	0.900 / 0.384	0.881 / 0.343	
QTAIM	C	2	0.762 / 0.278	0.777 / 0.293	84 / 85
	N	-3	0.728 / 0.021	0.692 / 0.022	
	Fe	3	0.848 / 0.344	0.827 / 0.305	
TFVC	C	2	0.694 / 0.235	0.709 / 0.245	85 / 86
	N	-3	0.701 / 0.020	0.665 / 0.021	

high-valent species is farther from the actual electron distribution than for the low-valent ones. This is actually expected because the "charges" associated to the OS are imaginary and the higher the oxidation state the more the actual atomic populations deviate from the formal ones.

It is worth to analyze in deeper detail the high-spin $[Fe(Pytacn)O(OH)]^{2+}$ species. The occupations of the last occupied and first unoccupied fragment orbitals, as well as the fragment-condensed charges and spin densities are collected on table (9.7). A threshold on the effective fragment occupations of 0.1 was used.

The EOS analysis yields a high-valent Fe^{V} species but the confidence indices for the alpha ($C_{\alpha}(\%)$=0.55) and beta ($C_{\beta}(\%)$=0.50) contributions are rather low, indicating a relevant Fe^{IV} character. Nevertheless, we find promising that

the EOS analysis recognizes the high-valent high-spin d^3 iron species as Fe^V, taking into account the observed values of the charge and spin densities. The ideal high-spin $Fe^V{=}O^{(2-)}$ picture is shown in fig. (9.4).

Table 9.7: EOS analysis and fragment-condensed charges and spin populations for the high-spin $[Fe(Pytacn)O(OH)]^{2+}$ species.

Method	Fragment	EOS	$\lambda^{\alpha}_{LO}/\lambda^{\alpha}_{FU}$	$\lambda^{\beta}_{LO}/\lambda^{\beta}_{FU}$	Charge	Spin density
UB3LYP	Fe	5	0.94 / 0.50	0.99 / 0.46	1.50	2.10
	OH	-1	0.69 / < 0.1	0.77 / < 0.1	0.11	0.12
	O	-2	0.64 / < 0.1	0.46 / < 0.1	-0.89	0.85
	Pytacn	0	0.55 / < 0.1	0.71 / < 0.1	1.28	-0.09

The expected spin density on the Fe atom should be close to 3 and that of the oxo moiety close to zero, as anticipated for a closed-shell O^{2-} species. The values reported in the table (9.7) are quite far from the ideal picture. The spin density on the Fe atom amounts to only 2.10 , and a significant value of 0.85 is found for the oxo moiety.

We tested the effect of using a different functional. In particular we explored the effect of the amount of exact exchange in it. It is often stated that the inclusion of exact exchange favors electron delocalization and spin-polarization. We carried out the analysis for Becke's half-and-half LYP and PBE functionals. The former includes 50% of exact exchange and the latter none. (B3LYP uses 20%.) As anticipated, when decreasing the amount of exact-exchange with the PBE functional the spin density on the oxo moiety slightly decreases (0.80), and when performing the EOS analysis one obtains a more clear cut Fe^V character, with $C_\alpha(\%)$=0.57 and $C_\beta(\%)$=0.60. Yet, the spin density on the Fe atom (1.99) is still much lower than in the ideal picture. Conversely, when increasing the exact exchange with BHandHLYP the spin density on the oxo moiety increases to 1.13 and the EOS analysis does not yield the Fe^V anymore.

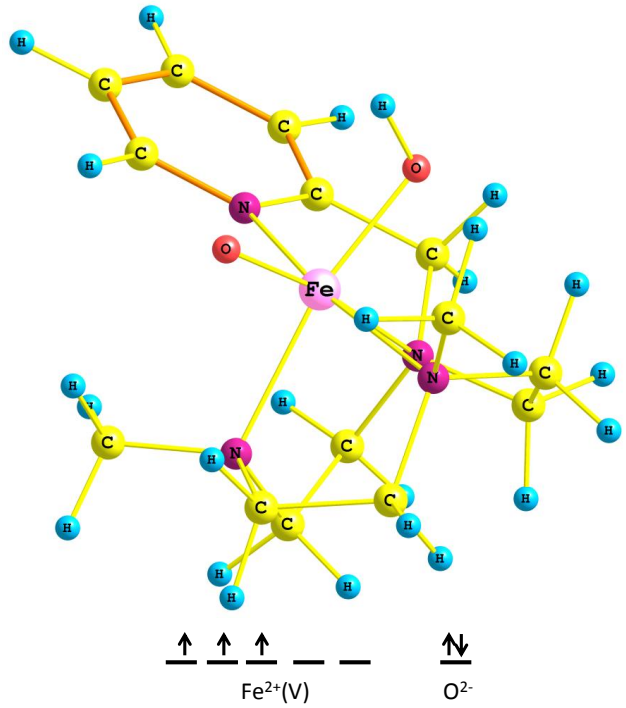

Figure 9.9: Ideal atomic orbital occupations for a high-spin $Fe^{V}{=}O^{(2-)}$ electron distribution.

9.5 The elusive $[Cu_3S_2]^{3+}$ core: a case of study

The trinuclear symmetric complex $[(tmedaCu)_3S_2]^{3+}$ [110] is depicted in fig. (9.10). The central $[Cu_3S_2]^{3+}$ core represents a model of the active site of metalloenzymes involved in biological processes. In recent years, its electronic structure has been the subject of vigorous debate in the literature.[110–114] The complex exist as an open-shell triplet , and discussions concerned the question whether there is a direct S-S bond in the $[Cu_3S_2]^{3+}$ core alongside with the formal OS of the Cu atoms. The conclusions of the earlier studies can be summarized as follows. According to Alvarez,[110] the $[Cu_3S_2]^{3+}$ core contains uncoupled S^{2-} ions compatible with a mixed-valence $[(Cu^{II})_2Cu^{III}]^{7+}$ configuration. Mealli and Hoffmann[112] suggested that the preferred configuration of the $[Cu_3S_2]^{3+}$ core is $[(Cu^{II})_2Cu^{I}]^{5+}$ with a bonded S_2^{2-} central unit. Berry[113] argued that the complex is best characterized as a noninnocent S_2^{3-} ligand antiferromagnetically coupled to three equivalent $[Cu^{II}]$ centers.

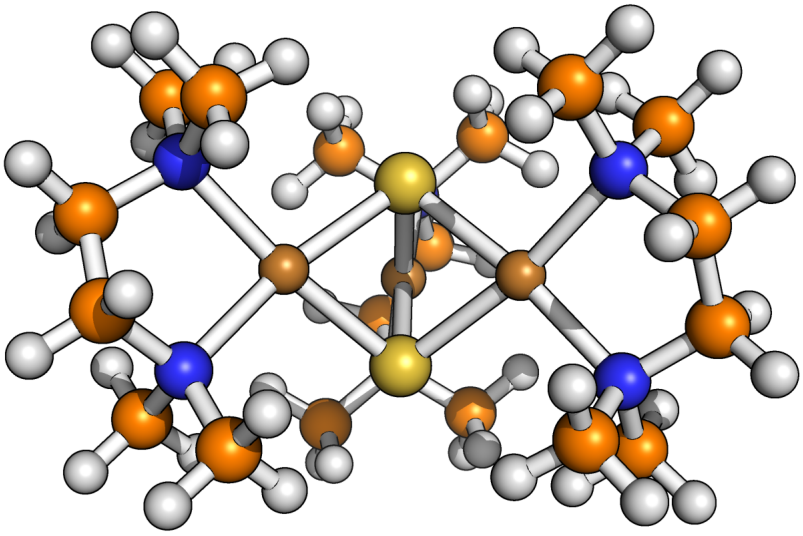

Figure 9.10: Schematic representation of the $[(\mathrm{tmedaCu})_3\mathrm{S}_2]^{3+}$ complex.

$$[(\mathrm{Cu^{II}})_2\mathrm{Cu^{III}}]^{7+}(\mathrm{S}^{2-})_2 \Longleftrightarrow [\mathrm{Cu_3^{II}}]^{6+}\mathrm{S}_2^{3-} \Longleftrightarrow [(\mathrm{Cu^{II}})_2\mathrm{Cu^{I}}]^{5+}\mathrm{S}_2^{2-}$$

Alvarez Berry Mealli,Hoffmann

In order to elucidate the nature of the S-S bonding we have reported a detailed analysis of the domain average Fermi holes (DAFH). To study the spin-spin interactions among the atoms in the $[\mathrm{Cu}_3\mathrm{S}_2]^{3+}$ core we have applied the local spin analysis. The final issue that we have addressed in this study is the formal oxidation state of the atoms in the $[\mathrm{Cu}_3\mathrm{S}_2]^{3+}$ core.

The DAFH analysis has provided a clear evidence of the presence of a weak 2c-3e S-S bond. The analysis of the local spin for this system shows rather significant local spin values on both Cu and S atoms, but extremely weak diatomic spin components among the atoms of the $[\mathrm{Cu}_3\mathrm{S}_2]^{3+}$ core, consistent with a delocalized spin picture. The local spin on Cu and S atoms is 0.23 and 0.25, respectively. No significant local spin is found on the N atoms of the ligands, even thought a small integrated spin density was found. The largest diatomic contribution is found between the Cu-Cu centers (0.023), whereas the value between Cu and S is just 0.005. Both the magnitude and the sign of

this interaction is not consistent with an antiferromagnetic coupling between the Cu centers and the S_2 unit. We have recently studied the same system at the CASSCF level of theory including 8 electrons in 8 orbitals.[114,115] At this level of theory the local spins on the atoms of the core ($\langle \hat{S}^2 \rangle_{\mathrm{Cu}} = 0.63$ and $\langle \hat{S}^2 \rangle_{\mathrm{S}} = 0.33$) are significantly larger specially for Cu. The diatomic terms ($\langle \hat{S}^2 \rangle_{\mathrm{CuS}} = -0.14$) reveal the presence of some antiferromagnetic coupling between the Cu and the S atoms (unpublished results).

The eff-AO analysis of the Cu and S atoms yielded 39 alpha and 36 beta atomic hybrid orbitals with occupation number larger than 0.98. These correspond to the full inner $1s^22s^22p^63s^23p^6$ shell and nine completely filled d-type atomic spin orbitals of each Cu atom and to the $1s^22s^22p^63s^2$ atomic orbitals of the S atoms. This electron distribution is consistent with the predominant Cu^{II} configuration of each of the three metal ions. However, the existence of non-negligible populated beta $d_{x^2-y^2}$ orbitals on the Cu atoms suggests a minor contamination of the predominant configuration by an admixture of contributions of delocalized $[Cu^{II}Cu^{II}Cu^{I}]$ configuration. The results of the EOS analysis for this system confirm that the principal configuration is the delocalized $[Cu^{II}Cu^{II}Cu^{II}]$ with $C_\alpha(\%)=0.61$ and $C_\beta(\%)=0.54$ (unpublished results).

By the time we were completing our study, a combined experimental and computational study by Sarangi *et al.*[114] concluded that the bonding in the $[Cu_3S_2]^{3+}$ core is delocalized. They suggested a weak S-S bonding interaction and also the existence of three equivalent copper centers close to but slightly more oxidized than Cu^{II}. Sarangi *et al.* used both DFT and CASSCF methods to describe the complex. However, the tentative assignation of oxidation states was performed by comparing the shape and intensity of the K-edge XAS spectra of the complex with that of "reference" complexes, not by any population analysis. For the Cu K-edge XAS spectra, the results were compared with the spectra of Cu(I), Cu(II) and Cu(III) complexes. For the S spectra dicopper(II) disulfido (S_2^{2-}) and supersulfido (S_2^{-}) complexes were used as reference. Our results agree with this picture except for the very last point (we find three equivalent copper centers close to but slightly *less* oxidized than Cu^{II}).

Chapter 10

Conclusions

The main conclusions drawn from this thesis are collected in this chapter. A new topological definition of fuzzy Voronoi cells is proposed, reproducing to some extent the results of the QTAIM analysis. It has been shown that the TFVC method represents a fast and simple atoms in molecules scheme that can be routinely used to extract chemical information from large-scale *ab initio* calculations. The method presented in this thesis can be regarded as a general-purpose computationally more efficient alternative to Bader's QTAIM.

We have derived a new general formulation based on the decomposition of the expectation value of the spin squared operator into atomic and diatomic contributions. It has been shown that there is a *continuum* of different formulations for the decomposition of $\langle\, \hat{S}^2 \,\rangle$ that fulfill a number of physical requirements.We have introduced an additional new criterion based upon the behaviour of single-electron systems that has permitted to derive a unique new general formulation applicable for both single-determinant and correlated wave functions. The scheme has been realized in the three-dimensional physical space and implemented for Bader's QTAIM and a number of fuzzy atom definitions. Very small local-spin values are obtained for genuine diamagnetic molecules treated at the correlated level, in conformity with the physical expectations.

Appropriate formulations have been derived for the so-called Hilbert-space

analysis. We have shown that there is an ambiguity affecting Mulliken-type decompositions in the selection of the indexes where the atoms are centered. We definitely recommend to adopt the assignment of subscripts that is consistent with Mulliken population analysis and the original Mayer-Wiberg bond orders. The results obtained are in good agreement with physical expectation and, in general, do not depend too much on the basis set. Nevertheless, for the particular case of acetylene one can obtain spurious results when performing the decomposition of quantities that explicitly depend upon the second-order density matrix, even with small basis sets. Löwdin's scheme is more robust and reliable in all cases.

We have illustrated in detail how the local spin analysis can provide valuable quantitative information about the presence of local spins and how these local spins couple. The connection between the local spin analysis and the nature of the chemical bond has been establish. The existence of significant local spins in molecules in pure singlet states is a signal for deviation from the classical bonding prototypes, with the C_2 molecule as a notorious example. The local spin analysis reveals that this molecule or any of its isoelectronic heterodiatomics is best described as a diradicaloid. Whether C_2 has a fourth bond or not may be a matter of semantics. What emerges from the local spin analysis is that the bonding here operates in a very different way compared to the prototypical one.

The local spin analysis can be used to unravel the electronic structure of diradical and triradical systems. The values of the atomic spin contributions for actual molecular systems can be used to define a measure of their k-radical character, $\Delta^{(k)}$. The smaller the $\Delta^{(k)}$ the closer the system is to an ideal system of k perfectly localized spin centers. Simple models provide the necessary reference atomic and diatomic contributions to $\langle \hat{S}^2 \rangle$ for perfectly localized spins. $\Delta^{(0)}$ and $\Delta^{(2)}$ are used to characterize diradicals and diradicaloids, whereas $\Delta^{(3)}$ is used for triradicals.

The magnitude and sign of the diatomic spin terms provides a picture of the entanglement between the local spin carriers. In the case of low-spin triradicals, the local spin analysis distinguishes between two patterns of spin distribution,

namely a singlet diradical uncoupled to an additional isolated radical center, and another one where the three spin centers exhibit interactions with each other. We have found that for a number of all σ, all π and σ–π organic triradicals the former picture is associated to smaller triplet stabilization energies. The larger the $\langle \hat{S}^2 \rangle_A$ value of a spin center is in a triradical, the easier it is to pull apart this particular radical center from the remaining two.

Another tool that aims at the characterization of the state of the atom within the molecule is Mayer's effective atomic orbitals. In this thesis the effective atomic orbitals have been realized for the first time in the framework of Bader's QTAIM. The eff-AOs form an orthonormalized set of numerical atomic orbitals, with occupation numbers that sum up to the respective QTAIM atomic populations. Only a limited number of eff-AOs exhibit significant occupation numbers, those associated with the core and valence shells of the atom. In the case of hypervalent atoms additional hybrids appear with small but non-negligible occupation numbers.

We have also shown that the MOs can be exactly (or up to a good accuracy, using a SVD procedure) expressed in terms of a much reduced set of orthogonal atomic basis functions. The Mulliken population analysis carried out on the basis of eff-AOs exactly reproduces the original QTAIM atomic populations of the atoms, indicating that the flaws of the classical Mulliken populations are rooted in the use of unsuitable atomic basis functions, not in its mathematical framework.

The eff-AOs can be utilized to derive the most appropriate atomic electron configurations for the atoms or molecular fragments in the molecule. A simple and general method to compute oxidation states from the analysis of the wave function has been introduced. The method provides effective oxidation states for all atoms and an overall confidence index that quantifies how reliable the formal assignation is. It can be applied in equal footing for any level of theory and for any chemical system except for hydrocarbons and extremely apolar species where the mere concept of oxidation state is less meaningful. For large systems a hierarchical strategy in which the system is first partitioned into

fragments/ligands appears more appropriate.

Finally, we have shed light into some of the pending issues concerning the electronic structure of the $[Cu_3S_2]^{3+}$ core. The DAFH analysis has provided a clear evidence of weak direct S-S interaction that can best be characterized as 2c-3e bond. The local spin analysis rules out the suggested antiferromagnetic coupling among the Cu atoms and S_2^{3-} unit, and rather confirms a delocalized-spin picture. The scrutiny of the effective atomic orbitals and their occupations points to an electron configuration of the Cu atoms close to a Cu^{II} formal oxidation state, but with a small admixture of a less oxidized Cu^{I} character.

Bibliography

1 Head-Gordon, M. *J. Phys. Chem.* **1996**, *100*, 13213–13225.

2 Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. *J. Phys. Chem. A* **2005**, *109*, 3957–3959.

3 Jansen, M.; Wedig, U. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 10026–10029.

4 McWeeny, R. *Rev. Mod. Phys.* **1960**, *32*, 335–369.

5 Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397–5403.

6 Ruedenberg, K. *Rev. Mod. Phys.* **1962**, *34*, 326–376.

7 Matito, E.; Feixas, F. DMN program. 2009; University of Girona (Spain) and University of Szczecin (Poland).

8 Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C.02. Gaussian, Inc., Pittsburgh, PA, 2003.

9 Pauncz, R. *Spin eigenfunctions*; Springer, 1979.

10 Szabo, A.; Ostlund, N. S. *Modern quantum chemistry: introduction to advanced electronic structure theory*; Courier Dover Publications, 2012.

11 Lewis, G. N. *J. Am. Chem. Soc.* **1916**, *38*, 762–786.

12 Giambiagi, M.; Giambiagi, M.; Jorge, F. *Theor. Chim. Acta (Berlin)* **1985**, *68*, 337–341.

13 Ángyán, J.; Rosta, E.; Surján, P. *Chem. Phys. Lett.* **1999**, *299*, 1–8.

14 Mayer, I. *Int. J. Quant. Chem.* **1986**, *29*, 477–483.

15 Bader, R.; Stephens, M. *Chem. Phys. Lett.* **1974**, *26*, 445–449.

16 Bader, R.; Stephens, M. *J. Am. Chem. Soc.* **1975**, *97*, 7391–7399.

17 Bochicchio, R.; Lain, L.; Torre, A.; Ponec, R. *J. Math. Chem.* **2000**, *28*, 83–90.

18 Angyan, J. G.; Loos, M.; Mayer, I. *J. Phys. Chem.* **1994**, *98*, 5244–5248.

19 Francisco, E.; Martín Pendás, A.; Blanco, M. A. *J. Chem. Theory Comput.* **2006**, *2*, 90–102.

20 Salvador, P.; Duran, M.; Mayer, I. *J. Chem. Phys.* **2001**, *115*, 1153–1157.

21 Poater, J.; Fradera, X.; Duran, M.; Solà, M. *Chem. Eur. J.* **2003**, *9*, 400–406.

22 Heyndrickx, W.; Salvador, P.; Bultinck, P.; Solà, M.; Matito, E. *J. Comput. Chem.* **2011**, *32*, 386–395.

23 Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833–1840.

24 Löwdin, P.-O. *Adv. Quantum. Chem.* **1970**, *5*, 185.

25 Hirshfeld, F. L. *Theor. Chim. Acta (Berlin)* **1977**, *44*, 129–138.

26 Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford Univ. Press: Oxford, 1990.

27 Mayer, I.; Salvador, P. *Chem. Phys. Lett.* **2004**, *383*, 368–375.

28 Matito, E.; Solà, M.; Salvador, P.; Duran, M. *Faraday Discuss.* **2007**, *135*, 325–345.

29 Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbó-Dorca, R. *J. Chem. Phys.* **2007**, *126*, 144111.

30 Lillestolen, T. C.; Wheatley, R. J. *Chem. Commun.* **2008**, 5909.

31 Bultinck, P.; Cooper, D. L.; Neck, D. V. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3424–3429.

32 Lillestolen, T. C.; Wheatley, R. J. *J. Chem. Phys.* **2009**, *131*, 144101.

33 Baker, J. *Theor. Chim. Acta (Berlin)* **1985**, *68*, 221–229.

34 Fonseca Guerra, C.; Handgraaf, J.-W.; Baerends, E. J.; Bickelhaupt, F. M. *J. Comput. Chem.* **2004**, *25*, 189–210.

35 Matito, E.; Poater, J.; Solà, M.; Duran, M.; Salvador, P. *J. Phys. Chem. A* **2005**, *109*, 9904–9910.

36 Salvador, P.; Mayer, I. *The Journal of chemical physics* **2004**, *120*, 5046.

37 Mayer, I.; Salvador, P. *J. Chem. Phys.* **2009**, *130*, 234106.

38 Becke, A. *J. Chem. Phys.* **1988**, *88*, 2547–2553.

39 Slater, J. C. *J. Chem. Phys.* **1964**, *41*, 3199.

40 Vyboishchikov, S. F.; Salvador, P.; Duran, M. *J. Chem. Phys.* **2005**, *122*, 244110.

41 Clark, A.; Davidson, E. *J. Chem. Phys.* **2001**, *115*, 7382–7392.

42 Clark, A.; Davidson, E. *Molec. Phys.* **2002**, *100*, 373–383.

43 Clark, A.; Davidson, E. *J. Phys. Chem. A* **2002**, *106*, 6890–6896.

44 Mayer, I. *Chem. Phys. Lett.* **2007**, *440*, 357–359.

45 Alcoba, D.; Lain, L.; Torre, A.; Bochicchio, R. *Chem. Phys. Lett.* **2009**, *470*, 136–139.

46 Mayer, I. *Chem. Phys. Lett.* **2009**, *478*, 323–326.

47 Mayer, I.; Matito, E. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11308–11314.

48 Alcoba, D.; Torre, A.; Lain, L.; Bochicchio, R. *Chem. Phys. Lett.* **2011**, *504*, 11308–11314.

49 Alcoba, D. R.; Torre, A.; Lain, L.; Bochicchio, R. C. *J. Chem. Theory Comput.* **2011**, *7*, 3560–3566.

50 Reiher, M. *Faraday Discuss.* **2006**, *135*, 97–124.

51 Herrmann, C.; Reiher, M.; Hess, B. *J. Chem. Phys.* **2005**, *122*, 034102.

52 Luzanov, A. V.; Prezhdo, O. V. *Molec. Phys.* **2007**, *105*, 2879–2891.

53 Podewitz, M.; Herrmann, C.; Malassa, A.; Westerhausen, M.; Reiher, M. *Chem. Phys. Lett.* **2008**, *451*, 301 – 308.

54 Podewitz, M.; Reiher, M. *Advances in Inorganic Chemistry* **2010**, *62*, 177–230.

55 Takatsuka, K.; Fueno, T.; Yamaguchi, K. *Theor. Chim. Acta (Berlin)* **1978**, *48*, 175–183.

56 Kutzelnigg, W.; Mukherjee, D. *J. Chem. Phys.* **2002**, *116*, 4787–4801.

57 Lain, L.; Torre, A.; Bochicchio, R. *J. Chem. Phys.* **2002**, *117*, 5497–5498.

58 Mayer, I. *Chem. Phys. Lett.* **1995**, *242*, 499–506.

59 Mayer, I. *J. Phys. Chem.* **1996**, *100*, 6249–6257.

60 Mayer, I. *Can. J. Chem.* **1996**, *74*, 939–942.

61 Mayer, I.; Bakó, I.; Stirling, A. *J. Phys. Chem. A* **2011**, *115*, 12733–12737.

62 Ponec, R. *J. Math. Chem.* **1997**, *21*, 323–333.

63 Ponec, R. *J. Math. Chem.* **1998**, *23*, 85–103.

64 Ponec, R.; Duben, A. J. *J. Comput. Chem.* **1999**, *20*, 760–771.

65 Ponec, R.; Yuzhakov, G.; Cooper, D. L. *Theor. Chim. Acta (Berlin)* **2004**, *112*, 419–430.

66 Ponec, R.; Yuzhakov, G.; Gironés, X.; Frenking, G. *Organometallics* **2004**, *23*, 1790–1796.

67 Ponec, R.; Yuzhakov, G. *Theor. Chim. Acta (Berlin)* **2007**, *118*, 791–797.

68 Ponec, R.; Cooper, D. L.; Savin, A. *Chem. Eur. J.* **2008**, *14*, 3338–3345.

69 Ponec, R.; Lendvay, G.; Chaves, J. *J. Comput. Chem.* **2008**, *29*, 1387–1398.

70 Ponec, R.; Lendvay, G.; Sundberg, M. R. *J. Phys. Chem. A* **2008**, *112*, 9936–9945.

71 Ponec, R.; Feixas, F. *J. Phys. Chem. A* **2009**, *113*, 8394–8400.

72 Ponec, R.; Feixas, F. *J. Phys. Chem. A* **2009**, *113*, 5773–5779.

73 Feixas, F.; Ponec, R.; Fiser, J.; Roithová, J.; Schröder, D.; Price, S. D. *The Journal of Physical Chemistry A* **2010**, *114*, 6681–6688.

74 Ponec, R.; Cooper, D. L. *Faraday Discuss.* **2007**, *135*, 31–42.

75 Cooper, D. L.; Ponec, R. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1319–1329.

76 Cioslowski, J. *Int. J. Quant. Chem.* **1990**, *38*, 015–028.

77 McNaught, A. D.; Wilkinson, A. *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*; Blackwell Scientific Publications, Oxford (1997), 1997.

78 Jørgensen, C. K. *Oxidation Numbers and Oxidation States*; Springer, 1969.

79 Chaudhuri, P.; Verani, C. N.; Bill, E.; Bothe, E.; Weyhermüller, T.; Wieghardt, K. *J. Am. Chem. Soc.* **2001**, *123*, 2213–2223.

80 Orpen, A. G. *Acta Crystallographica Section B* **2002**, *58*, 398–406.

81 Palenik, G. J. *Inorg. Chem.* **1997**, *36*, 3394–3397.

82 Brown, I. D.; Altermatt, D. *Acta Crystallographica Section B* **1985**, *41*, 244–247.

83 Thom, A. J. W.; Sundstrom, E. J.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 11297–11304.

84 Sit, P. H.-L.; Zipoli, F.; Chen, J.; Car, R.; Cohen, M. H.; Selloni, A. *Chem. Eur. J.* **2011**, *17*, 12136–12143.

85 Sit, P. H.-L.; Car, R.; Cohen, M. H.; Selloni, A. *Inorg. Chem.* **2011**, *50*, 10259–10267.

86 Knizia, G. *J. Chem. Theory Comput.* **2013**,

87 Jiang, L.; Levchenko, S. V.; Rappe, A. M. *Phys. Rev. Lett.* **2012**, *108*, 166403.

88 Aullón, G.; Alvarez, S. *Theor. Chim. Acta (Berlin)* **2009**, *123*, 67–73.

89 Salvador, P.; Ramos-Cordoba, E. APOST-3D program. 2012; Universitat de Girona (Spain).

90 Salvador, P.; Mayer, I. AFUZZY program. 2004; Universitat de Girona (Spain).

91 Salvador, P.; Mayer, I. ENPART program. 2005; Universitat de Girona (Spain).

92 Rodríguez, J. I.; K oster, A. M.; Ayers, P. W.; Santos-Valle, A.; Vela, A.; Merino, G. *J. Comput. Chem.* **2009**, *30*, 1082–1092.

93 Su, P.; Wu, J.; Gu, J.; Wu, W.; Shaik, S.; Hiberty, P. C. *J. Chem. Theory Comput.* **2011**, *7*, 121–130.

94 Shaik, S.; Danovich, D.; Wu, W.; Su, P.; Rzepa, H. S.; Hiberty, P. C. *Nat. Chem.* **2012**, *4*, 195–200.

95 Shaik, S.; Rzepa, H. S.; Hoffmann, R. *Angew. Chem. Int. Ed. Engl.* **2013**, *52*, 3020–3033.

96 Wenthold, P. G.; Squires, R. R.; Lineberger, W. *J. Am. Chem. Soc.* **1998**, *120*, 5279–5290.

97 De Proft, F.; von Ragué Schleyer, P.; van Lenthe, J. H.; Stahl, F.; Geerlings, P. *Chem. Eur. J.* **2002**, *8*, 3402–3410.

98 Yoshizawa, K.; Hatanaka, M.; Matsuzaki, Y.; Tanaka, K.; Yamabe, T. *J. Chem. Phys.* **1994**, *100*, 4453–4458.

99 C. R. Kemnitz, R. S. S.; Borden, W. T. *J. Am. Chem. Soc.* **1997**, *119*, 6564–6574.

100 Nguyen, H. M. T.; Dutta, A.; Morokuma, K.; Nguyen, M. T. *J. Chem. Phys.* **2005**, *122*, 154308.

101 Höltzl, T.; Veszprémi, T.; Nguyen, M. T. *Chem. Phys. Lett.* **2010**, *499*, 26–30.

102 Neuhaus, P.; Sander, W. *Angew. Chem. Int. Ed. Engl.* **2010**, *122*, 7435–7438.

103 Krylov, A. *J. Phys. Chem. A* **2005**, *109*, 10638–10645.

104 Cristian, A.-M. C.; Shao, Y.; Krylov, A. I. *J. Phys. Chem. A* **2004**, *108*, 6581–6588.

105 Nguyen, H. M. T.; Höltzl, T.; Gopakumar, G.; Veszprémi, T.; Peeters, J.; Nguyen, M. T. *Chem. Phys.* **2005**, *316*, 125–140.

106 Slipchenko, L. V.; Krylov, A. I. *J. Chem. Phys.* **2003**, *118*, 9614–9622.

107 Slipchenko, L. V.; Munsch, T. E.; Wenthold, P. G.; Krylov, A. I. *Angew. Chem. Int. Ed. Engl.* **2004**, *116*, 760–763.

108 Wang, T.; Krylov, A. I. *Chem. Phys. Lett.* **2006**, *425*, 196–200.

109 Fillol, J. L.; Codolà, Z.; Garcia-Bosch, I.; Gómez, L.; Pla, J. J.; Costas, M. *Nat Chem* **2011**, *3*, 807–813.

110 Brown, E. C.; York, J. T.; Antholine, W. E.; Ruiz, E.; Alvarez, S.; Tolman, W. B. *J. Am. Chem. Soc.* **2005**, *127*, 13752–13753.

111 Alvarez, S.; Hoffmann, R.; Mealli, C. *Chem. Eur. J.* **2009**, *15*, 8358–8373.

112 Mealli, C.; Ienco, A.; Poduska, A.; Hoffmann, R. *Angew. Chem. Int. Ed. Engl.* **2008**, *47*, 2864–2868.

113 Berry, J. F. *Chem. Eur. J.* **2010**, *16*, 2719–2724.

114 Sarangi, R.; Yang, L.; Winikoff, S. G.; Gagliardi, L.; Cramer, C. J.; Tolman, W. B.; Solomon, E. I. *J. Am. Chem. Soc.* **2011**, *133*, 17180–17191.

115 Carrasco, R.; Aullón, G.; Alvarez, S. *Chem. Eur. J.* **2009**, *15*, 536–546.