Theor Chem Acc (2011) 128:433–444
DOI 10.1007/s00214-010-0809-4

REGULAR ARTICLE

Generalized electron number distribution functions: real space versus orbital space descriptions

E. Francisco · A. Martín Pendás · Miguel A. Blanco

Received: 25 June 2010 / Accepted: 27 August 2010 / Published online: 19 September 2010

Abstract The algebraic expressions previously derived to compute the electron number distribution functions (EDF) for exhaustive partitions of the physical space into sharp-boundary atoms are generalized to allow for the use of fuzzy atoms and orbital-based partitions. In some of the latter, the atomic overlap matrix required to obtain the EDF is analytical. This makes them attractive alternatives to other definitions, as the one based on the atomic basins of the quantum theory of atoms in molecules (QTAIM), which are more physically sound but also much more demanding computationally. We will compute the EDF for a series of test molecules using different fuzzy and orbital-based partitions and compare them to QTAIM EDF. The effects of electron correlation and the use of the core approximation on the EDF will also be explored.

Keywords Real space partition · Electron number distribution function · Atomic overlap matrix · Quantum theory of atoms in molecules · Delocalization index

1 Introduction

According to the basic postulates of Quantum Mechanics, the multi-electron wave function Ψ contains all the information that is possible to obtain from a molecular system. However, given that the Hamiltonian contains only one- and two-particle terms, the complete Ψ is usually not required in most applications of Quantum Chemistry. Instead, the two-electron reduced density matrix (2-RDM) [1],

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

where N is the total number of electrons, and $\boldsymbol{x} = \{\boldsymbol{r},\sigma\}$ denotes the combined Cartesian–spin electron coordinates, is generally enough to extract from it the practical totality of *numbers* that chemists need to discuss molecular properties and concepts. By integrating $\rho_2(\boldsymbol{x}_1,\boldsymbol{x}_2)$ over $\boldsymbol{r}_2,\sigma_1$, and σ_2, the total electron density, $\rho(\boldsymbol{r})$, which is the key function of the Quantum Theory of Atoms in Molecules (QTAIM), is obtained [2]. The 2-RDM is a particular case of more general mathematical functions known as coarse-grained density matrices (CGDM) [3–5] obtained by integrating $|\Psi(1,N)|^2$ over arbitrary domains of the physical 3D space for only a subset of the N electrons. Another particular CGDM is the multivariate electron number distribution function (EDF) [6–11]. The latter arises when one tries to solve the following statistical question: Assuming that the 3D space is partitioned into m arbitrary but disjoint domains $\Omega_1,\Omega_2,\ldots,\Omega_m$, which is the probability of finding exactly and simultaneously n_1 out of the N electrons within domain Ω_1, n_2 electrons within domain $\Omega_2,\ldots$, and n_m electrons $(n_1 + n_2 + \cdots + n_m = N)$ within basin Ω_m, where all $n_i \in \mathbb{N}$? Several key concepts of chemical bonding theory, like electron pair sharing [12–18], multicenter bonding [19, 20, 21, 22], polarity of bonds, charge transfer, the origin of the additivity of the properties of functional groups, etc. [10, 11, 23–25], are appropriately discussed or more easily understood in terms of the EDF,

---

Published as part of the special issue celebrating theoretical and computational chemistry in Spain.

Miguel A. Blanco: deceased.

E. Francisco (✉) · A. M. Pendás · M. A. Blanco
Facultad de Química, Departamento de Química Física y Analítica, Universidad de Oviedo, 33006 Oviedo, Spain
e-mail: evelio@carbono.quimica.uniovi.es

A. M. Pendás
e-mail: angel@fluor.quimica.uniovi.es

whenever the domains or regions in which the 3D space is partitioned are endowed with a relevant chemical and physical meaning. Among the more physically sound partitions of the 3D space, the ones based on the loge theory of Daudel [26, 27], the atomic domains of the QTAIM [2], or the core, valence, and lone pair regions obtained from the use of the electron localization function (ELF) [28–30] have been particularly fruitful. Neither the EDF nor the more general CGDM can be obtained from $\rho_2(\boldsymbol{x}_1,\boldsymbol{x}_2)$ and need to be calculated directly from the complete $\Psi(1,N)$.

Our group has recently developed efficient methods to obtain the EDF based on QTAIM atomic basins, for single- and multi-determinant wave functions, and for an arbitrary number ($m \geq 2$) of regions [7–9], and used them to provide extra information in energy partition discussions [31–34], to classify chemical bonds [10], and provide a new view of the bond paths of the Quantum Theory of Atoms in Molecules (QTAIM) [2] as privileged quantum-mechanical exchange channels [24].

One of the main difficulties with EDF based on the QTAIM lies in the irregular forms of the QTAIM atomic basins, due to their sharp and well-defined boundaries. This makes computationally very costly to obtain the atomic overlap matrix (AOM) elements between the molecular orbitals (MO) (the main mathematical ingredients that enter the EDF expressions), as the inter-atomic surfaces have to be previously determined. This has been one of the reasons that has motivated us to extend the EDF concept to other partitions of the 3D space for which the AOM is more easily obtained. The other reason has been to establish a connection between the EDF expressions associated with different 3D space partitions and those of the homonym electron population analyses [35, 36], such as Mulliken [37], Löwdin [38], Becke [39], and Hirshfeld [40].

The rest of the article has been organized as follows. In Sect. 2, the expressions needed to obtain the EDF for exhaustive, fuzzy, and orbital-based partitions of the 3D space are worked out. How several relevant chemical bonding concepts are related to EDF is explained in Sect. 3. In Sect. 4, we describe the computational details and test systems for which the EDF have been computed. The obtained results are discussed and compared with each other in Sect. 5. Finally, the conclusions of this work are summarized in Sect. 6.

2 Electron number distribution functions in real and orbital spaces

Let a N-electron molecule be described by the wave function Ψ and an exhaustive partition of the 3D space into m disjoint domains $\Omega_k\left(\cup_{k=1}^{m}\Omega_k = \mathcal{R}^3, \Omega_k \cap \Omega_l = \emptyset, k \neq l\right)$. The probability that exactly n_1 electrons are within Ω_1, n_2 electrons within $\Omega_2, \ldots,$ and n_m electrons within $\Omega_m (0 \leq n_i \leq N, n_1 + n_2 + \cdots + n_m = N)$ is given by [26]

$$p(S) = N!\Lambda \int\limits_D |\Psi|^2 \mathrm{d}\boldsymbol{r}_1 \cdots \mathrm{d}\boldsymbol{r}_N, \tag{2}$$

where $S \equiv \{n_1, n_2,\ldots, n_m\}$, the summation over the spin variables $\sigma_i (i = 1,N)$ is implicitly assumed, $N!\Lambda = N!/(n_1!n_2! \cdots n_m!)$, and D is a multi-dimensional domain in which the first n_1 electrons are integrated over Ω_1, the second n_2 electrons over $\Omega_2, \ldots,$ and the last n_m electrons over Ω_m. There are $N_S = (N + m - 1)!/[N!(m - 1)!]$ possible distributions of the N electrons within the m domains, and the set $p(S)$ $(S = 1,\ldots, N_S)$ defines the EDF of the system for this specific space partition. To find an unified approach for both exhaustive and fuzzy partitions of the 3D space, it is useful to define $W_S = W_1 \times \cdots \times W_m$, where $W_1 = w_1(\boldsymbol{r}_1)\ldots w_1(\boldsymbol{r}_{n_1}), W_2 = w_2(\boldsymbol{r}_{n_1+1})\ldots w_2(\boldsymbol{r}_{n_1+n_2})$, etc, and $w_k(\boldsymbol{r})$ is the single-electron step function

$$w_k(\boldsymbol{r}) = \begin{cases} 1 & \text{if } \boldsymbol{r} \in \Omega_k \\ 0 & \text{elsewhere.} \end{cases} \tag{3}$$

Then, $p(S)$ may also be written such that the integration is over $\mathcal{R}^3$ for the N electrons:

$$p(S) = N!\Lambda \int W_S |\Psi|^2 \mathrm{d}\boldsymbol{r}_1 \cdots \mathrm{d}\boldsymbol{r}_N. \tag{4}$$

The above expression for $p(S)$ corresponds to an exhaustive partition of $\mathcal{R}^3$. However, it may also be applied to a so-called *fuzzy partition*. This is defined as one in which the domains $\Omega_1, \Omega_2, \ldots, \Omega_m$ are not mutually exclusive but separated by fuzzy boundaries. The transition from a domain to another one happens smoothly and not sharply, as in exhaustive partitions. Mathematically, a fuzzy partition may be achieved by introducing a set of weight functions $\omega_1(\boldsymbol{r}), \omega_2(\boldsymbol{r}), \ldots, \omega_m(\boldsymbol{r})$, associated respectively with $\Omega_1, \Omega_2, \ldots, \Omega_m$, such that, at every point $\boldsymbol{r}, 0 \leq \omega_k(\boldsymbol{r}) \leq 1$, and

$$\sum_{k=1}^{m} \omega_k(\boldsymbol{r}) = 1(\boldsymbol{r}). \tag{5}$$

The $\omega_k(\boldsymbol{r})$'s change continuously in a fuzzy partition, although one may expect that, provided that Ω_k represents an *atom in a molecule*, $\omega_k(\boldsymbol{r})$ is close to unity as one approaches to the corresponding nucleus. An exhaustive partition of $\mathcal{R}^3$, as for instance the QTAIM partition, is a limiting case of the above general framework in which only one of the m $\omega_k(\boldsymbol{r})$'s is different from zero (and thus equal to unity) at every $\boldsymbol{r}$.

In this work, we have considered two fuzzy partitions of $\mathcal{R}^3$: Becke's partition [39], based on the auxiliary geometric weight function $w_k(\boldsymbol{r})$ defined by this author to perform the integration over $\mathcal{R}^3$ of a function $F(\boldsymbol{r})$ by using $\int F(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = \int \sum_k w_k(\boldsymbol{r})F(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = \sum_k \int F_k(\boldsymbol{r})\mathrm{d}\boldsymbol{r}$, and Hirshfeld's or stockholder partition [40]. Becke's weights are defined by the equations:

$$w_k(\boldsymbol{r}) = P_k(\boldsymbol{r}) \Big/ \sum\nolimits_k P_k(\boldsymbol{r}) \tag{6}$$

$$P_k(\boldsymbol{r}) = \prod_{l \neq k} \frac{1}{2}\left[1 - \underbrace{h[h[\ldots h(v_{kl})]]}_{v\ times}\right], \tag{7}$$

$$h(v_{kl}) = \left(3v_{kl} - v_{kl}^3\right)/2, \tag{8}$$

$$v_{kl} = \mu_{kl} + a_{kl}\left(1 - \mu_{kl}^2\right), \tag{9}$$

$$\mu_{kl} = (r_k - r_l)/R_{kl}, \tag{10}$$

$$a_{kl} = [(R_l/R_k) - (R_k/R_l)]/4, \tag{11}$$

where r_k (r_l) is the distance to the nucleus of atom k (l), R_{kl} denotes the inter-nuclear distance between the nuclei of atoms k and l, and R_k and R_l are atomic size adjustable parameters. $w_k(\boldsymbol{r})$ is very close to 1.0 at the nucleus of atom k and decays to zero on approaching the nucleus of a different atom. The stiffness of the cutoff between different atoms may be enhanced by increasing the parameter v that gives the number of times that the $h(v_{kl})$ polynomial is iterated to obtain $P_k(\boldsymbol{r})$ in Eq. 7. In the limit $v \to \infty, \mathcal{R}^3$ is exhaustively partitioned into disjoint atomic regions. Here, we will use a topological-based Becke's partition: if there exists a bond critical point between atom k and l, R_k and R_l are taken as R_k^{top} and R_l^{top}, the topological radii [2] of atoms k and l, respectively; otherwise, R_k and R_l are the Bragg-Slater radii of atoms k and l [41].

Hirshfeld's $w_k(\boldsymbol{r})$'s are defined as [40]

$$w_k(\boldsymbol{r}) = \frac{\rho_k^0(\boldsymbol{r})}{\sum_k \rho_k^0(\boldsymbol{r})} = \frac{\rho_k^0(\boldsymbol{r})}{\rho^0(\boldsymbol{r})}, \tag{12}$$

where $\rho_k^0(\boldsymbol{r})$ is a reference atomic density for atom k. Usually, this is taken as the ground-state spherically averaged density of neutral atom k, and $\rho^0(\boldsymbol{r})$ coincides with the promolecular density, even though other definitions are possible [36]. An improved iterative Hirshfeld partition has recently been developed [42] that eliminates some of the inconsistencies of the standard one related to the Kullback-Leibler entropy deficiency functional [43]. However, the simplest version defined by Eq. 12 will be used here.

Before defining the EDF based on orbital space partitions, it is convenient to outline the algebraic method used to compute the EDF for a fuzzy or exhaustive partition of $\mathcal{R}^3$ (Further details have been given elsewhere [7, 9]). Let us assume that Ψ is a linear combination of M Slater determinants

$$\Psi = \sum_{r=1}^{M} \frac{d_r}{\sqrt{N!}} \left|\varphi_1^r(1)\varphi_2^r(2)\cdots\varphi_N^r(N)\right|, \tag{13}$$

where the φ_i^r's are assumed to be real and orthonormal. Squaring Ψ and substituting in Eq. 4, we obtain

$$p(S) = \sum_{r,s}^{M} \mathrm{d}_r \mathrm{d}_s p_{rs}(S), \quad \text{where} \tag{14}$$

$$p_{rs}(S) = \Lambda \int W_S \sum_{\boldsymbol{q} \in \mathbb{S}_N} D_{rs}(\boldsymbol{q})\mathrm{d}\boldsymbol{r}_1 \cdots \mathrm{d}\boldsymbol{r}_N, \tag{15}$$

$$D_{rs}(\boldsymbol{q}) = \begin{vmatrix} \varphi_1^r(q_1)\varphi_1^s(q_1) & \ldots & \varphi_1^r(q_N)\varphi_N^s(q_N) \\ \vdots & \ddots & \vdots \\ \varphi_N^r(q_1)\varphi_1^s(q_1) & \ldots & \varphi_N^r(q_N)\varphi_N^s(q_N) \end{vmatrix}, \tag{16}$$

$\mathbb{S}_N$ is the set of $N!$ permutations of $\{1, 2,\ldots, N\}$, and $\boldsymbol{q} = \{q_1, q_2, \ldots, q_N\}$ is one of these permutations. Carrying out the summation over the spin variables σ_i and the spatial integration in Eq. 15, we have

$$p_{rs}(S) = \Lambda \sum_{\{\boldsymbol{q}\} \in \mathbb{S}_N} \begin{vmatrix} \mathbf{S}_{11}^{rs}(q_1) & \ldots & \mathbf{S}_{1N}^{rs}(q_N) \\ \vdots & \ddots & \vdots \\ \mathbf{S}_{N1}^{rs}(q_1) & \ldots & \mathbf{S}_{NN}^{rs}(q_N) \end{vmatrix}, \tag{17}$$

where

$$\mathbf{S}_{ij}^{rs}(q_j) = \begin{cases} \left(\mathbf{S}_{ij}^{rs}\right)_{\Omega_1} & \text{for} \quad 1 \leq q_j \leq n_1 \\ \left(\mathbf{S}_{ij}^{rs}\right)_{\Omega_2} & \text{for} \quad n_1 + 1 \leq q_j \leq n_1 + n_2 \\ \quad \vdots & \end{cases} \tag{18}$$

and $\left(\mathbf{S}_{ij}^{rs}\right)_{\Omega_k}$ is given by

$$\left(\mathbf{S}_{ij}^{rs}\right)_{\Omega_k} = \delta\left(\sigma_i^r, \sigma_j^s\right) \int \omega_k(\boldsymbol{r})\varphi_i^r(\boldsymbol{r})\varphi_j^s(\boldsymbol{r})d\boldsymbol{r}. \tag{19}$$

In Eq. 19, $\delta\left(\sigma_i^r, \sigma_j^s\right) = 1(0)$ for φ_i^r and φ_j^s having equal (different) spins. When $\omega_k(\boldsymbol{r})$ in this equation corresponds to QTAIM partition (Eq. 3), $\left(\mathbf{S}_{ij}^{rs}\right)_{\Omega_k}$ becomes $\langle\varphi_i^r|\varphi_j^s\rangle_{\Omega_k} = \int_{\Omega_k} \varphi_i^r(\boldsymbol{r})\varphi_j^s(\boldsymbol{r})d\boldsymbol{r}$, an element of the AOM. It is in this sense that $\left(\mathbf{S}_{ij}^{rs}\right)_{\Omega_k}$ may be called a generalized AOM element. Due to Eq. 5, they satisfy

$$\sum_{k=1}^{m} \left(\mathbf{S}_{ij}^{rs}\right)_{\Omega_k} = \begin{cases} 1 & \text{for} \quad \varphi_i^r = \varphi_j^s \\ 0 & \text{for} \quad \varphi_i^r \neq \varphi_j^s. \end{cases} \tag{20}$$

The summation in the second member of Eq. 17 contains $N!$ terms, so that its computation is not possible except for small N's. However, the identity

$$\sum_S t_1^{n_1} t_2^{n_2} \cdots t_m^{n_m} p_{rs}(S) = \det\left[\sum_{k=1}^{m} t_k \mathbf{S}^{rs}_{\Omega_k}\right], \tag{21}$$

where $t_m = 1$ and $\mathbf{S}^{rs}_{\Omega_k}$ is the $(N \times N)$ matrix with $\left(\mathbf{S}^{rs}_{ij}\right)_{\Omega_k}$ elements given by Eq. 19 is valid for arbitrary $t_1, \ldots, t_{m-1}$ values [7]. Hence, after multiplying both members of Eq. 14 by $t_1^{n_1} \times \cdots \times t_m^{n_m}$, performing the summation over all possible resonance structures S, and using Eq. 21, the following linear system of N_S equations is obtained.

$$\sum_S t_1^{n_1} \cdots t_m^{n_m} p(S) = \sum_{r,s}^{M} d_r d_s \det\left[\sum_{k=1}^{m} t_k \mathbf{S}^{rs}_{\Omega_k}\right]. \tag{22}$$

To obtain the N_S unknowns $p(S)$, the following three steps are performed: (1) Assign N_S values to the set $\{t_1, t_2, \ldots, t_{m-1}\}$ (remember that $t_m = 1$), (2) for each of these N_S sets, compute the coefficients $t_1^{n_1} t_2^{n_2} \ldots t_m^{n_m}$ at the right-hand side of the linear system 22, and (3) solve the linear system.

The above method to determine the EDF is completely general. However, it can be formulated in a simpler form for single-determinant wave functions $(M = 1, d_r = 1, \Psi = (N!)^{-1/2}\det[\varphi_1(1)\varphi_2(2)\cdots\varphi_N(N)])$ and only two domains $(m = 2)\Omega$ and $\Omega' = \mathcal{R}^3 - \Omega$. In this case, $p_{rs}(S) \equiv p(S)$ and Eq. 17 transforms to [6]

$$p^{\Omega}(n) = \Lambda \sum_{\mathbf{q}\in\mathbb{S}_N} \det\left[\tilde{\mathbf{S}}_{ij}(q_j)\right], \tag{23}$$

where $p^{\Omega}(n)$ is the probability that n electrons are in Ω (and hence $N - n$ in Ω'), $\Lambda = [n!(N-n)!]^{-1}$, and $\tilde{\mathbf{S}}_{ij}(q_j) = \left(\mathbf{S}_{ij}\right)_{\Omega}$ or $\tilde{\mathbf{S}}_{ij}(q_j) = \left(\mathbf{S}_{ij}\right)_{\Omega'} = \delta_{ij} - \left(\mathbf{S}_{ij}\right)_{\Omega}$, depending on whether $(1 \le q_j \le n)$ or $(n < q_j \le N)$, respectively. On the other hand, Ψ is invariant under an unitary rotation of the φ_i's. In particular, we can choose a new set ϕ, related to the φ_i's by $\phi = \varphi\mathbf{U}$, where $\mathbf{U}$ is the unitary matrix that diagonalizes $\mathbf{S}_{\Omega}$; i.e. $\mathbf{S}_{\Omega}\mathbf{U} = \mathbf{U}\mathbf{s}$. In this new ϕ basis, $\mathbf{S}_{\Omega}(\phi)$ and $\mathbf{S}_{\Omega'}(\phi)$ become $\mathbf{s}$ and $\mathbf{1} - \mathbf{s}$, respectively, where $\mathbf{s}$ is a $N \times N$ diagonal matrix. Hence, $p^{\Omega}(n)$ transforms to

$$p^{\Omega}(n) = \Lambda \sum_{\mathbf{q}\in\mathbb{S}_N} \prod_{j=1}^{N} \lambda_j(q_j) \begin{cases} \lambda_j(q_j \le n) = s_j \\ \lambda_j(q_j > n) = 1 - s_j. \end{cases} \tag{24}$$

Since electrons 1 to n and $n + 1$ to N are associated with Ω and Ω', respectively, each of the $\Lambda^{-1} = n!(N-n)!$ restricted permutations of these two disjoint subsets give the same $\prod_{j=1}^{N} \lambda_j(q_j)$, canceling the overall Λ factor in Eq. 24. $p^{\Omega}(n)$ is thus given by the classical probability that a set of N independent electrons, with probabilities s_j and $1 - s_j$ $(j = 1, \ldots, N)$ to lie in Ω and Ω', respectively, is partitioned such that n of them are found in the first, and $N - n$ in the second domain. Besides this, $p^{\Omega}(n) = a_n^N$, where the a_k^i coefficients are evaluated using the following recursive formulas [6]

$$a_k^i = \begin{cases} a_0^i = (1 - s_i)a_0^{i-1} \\ a_k^i = s_i a_{k-1}^{i-1} + (1 - s_i)a_k^{i-1} \\ a_i^i = s_i a_{i-1}^{i-1} \end{cases} \tag{25}$$

with $a_0^0 = 1, 1 \le i \le N$, and $1 \le k \le i - 1$.

The computation of $p^{\Omega}(S)$ can be accelerated in several ways. First, we should notice that $\mathbf{S}_{\Omega}$, or each $\mathbf{S}^{rs}_{\Omega_k}$ in the case of multi-determinant wave functions, is blocked into α and β sub-matrices. This means that, for single-determinant wave functions and no matter whether the φ or ϕ basis is used, the α and β EDF, $p^{\Omega}_{\alpha}(n^{\alpha})$ and $p^{\Omega}_{\beta}(n^{\beta})$, can be independently calculated, and the spin-resolved probability that exactly n^{α} α electrons and n^{β} β electrons lie within Ω (and hence that $N^{\alpha} - n^{\alpha}$ α and $N^{\beta} - n^{\beta}$ β electrons are in Ω', where N^{σ} is the total number of electrons with spin σ) is obtained as

$$p^{\Omega}\left(n^{\alpha}; n^{\beta}\right) = p^{\Omega}_{\alpha}(n^{\alpha}) \otimes p^{\Omega}_{\beta}\left(n^{\beta}\right). \tag{26}$$

$p^{\Omega}(n)$ is obtained by adding all the $p^{\Omega}\left(n^{\alpha}; n^{\beta}\right)$ with $n^{\alpha} + n^{\beta} = n$. Needless to say, the α and β EDF are equal for closed-shell wave functions.

For multi-determinant wave functions, the α and β EDF are not to each other independent. However, an equation very similar to Eq. 26 can still be formulated for each $p_{rs}(S)$:

$$p_{rs}\left(S^{\alpha}; S^{\beta}\right) = p_{rs}(S^{\alpha}) \otimes p_{rs}\left(S^{\beta}\right), \tag{27}$$

where $S^{\sigma} = \left(n_1^{\sigma}, n_2^{\sigma}, \ldots, n_m^{\sigma}\right)$ defines a resonance structure for spin σ $(\sigma = \alpha, \beta)$. For each (r, s) pair, $p_{rs}(S^{\sigma})$ is computed by solving a linear system like Eq. 21 with S^{σ} instead of S, n_k^{σ} instead of n_k $(k = 1, \ldots, m)$, and the σ block of the generalized AOM instead of the full $\mathbf{S}^{rs}_{\Omega_k}$ matrix. The full spin-resolved EDF $p(S^{\alpha}; S^{\beta})$, that gives the probability that exactly $n_1^{\alpha}, n_2^{\alpha}, \ldots, n_m^{\alpha}$ α electrons are in $\Omega_1, \Omega_2, \ldots, \Omega_m$, respectively, and, simultaneously, $n_1^{\beta}, n_2^{\beta}, \ldots, n_m^{\beta}$ β electrons are in $\Omega_1, \Omega_2, \ldots, \Omega_m$, respectively, can then be obtained as

$$p\left(S^{\alpha}; S^{\beta}\right) = \sum_{r,s}^{M} d_r d_s p_{rs}\left(S^{\alpha}; S^{\beta}\right). \tag{28}$$

From it, $p(S)$ is computed by adding all the $p(S^{\alpha}; S^{\beta})$'s with $n_1^{\alpha} + n_1^{\beta} = n_1, n_2^{\alpha} + n_2^{\beta} = n_2, \ldots,$ and $n_m^{\alpha} + n_m^{\beta} = n_m$. Solving many times the linear system 21 for each spin instead of solving only once the linear system 22 is advantageous for different reasons: (a) The smaller dimension of the linear system 21 ($N_S^{\sigma} \times N_S^{\sigma}$, where $N_S^{\sigma} = (N^{\sigma} + m - 1)!/[N^{\sigma}!(m-1)!]$) with respect to that of linear system 22 ($N_S \times N_S$) makes computationally much cheaper to obtain the right-hand side of the former. Moreover, the double summation over r and s, present in Eq. 22, is absent from Eq. 21. A smaller dimension also contributes to minimize the numerical instabilities of this inherently ill-conditioned system; (b) The spin-resolved EDF $p(S^{\alpha}; S^{\beta})$ contains considerably more chemically relevant information than its spinless counter part $p(S)$; (c) Usually, a given configuration of α or β spin-orbitals appears repeated many times when the N-electron wave function $\Psi(1,N)$ is expanded in terms of Slater determinants. As a consequence, it is not necessary to solve Eq. 21 $M \times (M+1)/2$ times, but only $M^{\sigma} \times (M^{\sigma}+1)/2$, where M^{σ} is the number of strictly different configurations built in with σ spin-orbitals that appears in the expansion of $\Psi(1,N)$; and (d) For closed-shell molecules, the set of linear systems 21 for the α and β electrons is the same, and has thus to be solved only for one of the spin components.

N_S^{σ} grows very quickly with N^{σ} and m. Previously, we suggested a way to maintain N_S^{σ} as smaller as possible consisting in restricting the number of resonance structures by avoiding those that do assign an unphysical number of electrons to each domain Ω_k. However, we have observed that the computed probabilities are most times inconsistent (for instance, some probabilities are negative or higher than 1.0, the sum of probabilities differs from 1.0 by large, etc.) when this method is used. A more stable and useful method to avoid an excessively large N_S^{σ} is to use the core approximation: when a generalized diagonal AOM element, $(\mathbf{S}_{ii})_{\Omega_k}$, is greater than $S_{\max}$, where $S_{\max}$ is a number close to 1.0, the orbital is assumed to be fully localized in Ω_k, and consequently excluded from the EDF calculation. In single-determinant wave functions, the method works more efficiently if the starting orbitals φ_i [usually the canonical orbitals within the Hartree-Fock (HF) approximation or the Kohn-Sham orbitals in the Density Functional Theory (DFT)] are previously subjected to an isopycnic (unitary) localization that leaves $\Psi(1,N)$ unaltered [44].

Let us now introduce the orbital space version of the EDF concept. Assuming that each molecular orbital (MO) φ_i is a linear combination of basis functions χ_a,

$$\varphi_i = \sum_a \chi_a C_{ai}, \tag{29}$$

the EDF in an orbital space is defined in terms of generalized AOM elements $(\mathbf{S}_{ij})_{\Omega_k}$ given by

$$(\mathbf{S}_{ij})_{\Omega_k} = \sum_{a,b} C_{ai} C_{bj} \hat{\omega}_k^{ab} \Sigma_{ab}, \tag{30}$$

where Σ_{ab} are the overlap integrals between the basis functions, $\Sigma_{ab} = \langle \chi_a | \chi_b \rangle$, and $\hat{\omega}_k^{ab}$ is an unequivocally defined operator that acting on Σ_{ab} associates its full value to the atom k of the molecule (For simplicity, we will assume that each Ω_k is a single atom. If this is not the case, each $(\mathbf{S}_{ij})_{\Omega_k}$ can be obtained by adding up the corresponding AOM elements of all the atoms that belong to Ω_k). Choosing the $\hat{\omega}_k^{ab}$ such that $\sum_k \hat{\omega}_k^{ab} = \hat{1}$, the $(\mathbf{S}_{ij})_{\Omega_k}$'s satisfy

$$\sum_k (\mathbf{S}_{ij})_{\Omega_k} = \sum_k \sum_{a,b} C_{ai} C_{bj} \hat{\omega}_k^{ab} \Sigma_{ab} = \delta_{ij}. \tag{31}$$

In this work, we have used three different EDF definitions based on orbital space partitions: The minimally deformed atoms partition of Fernández Rico et al. [45, 46] (MinDef partition in the following), Mulliken's partition and Löwdin's partition. These names come from their close relationship to the corresponding homonym population analysis methods.

In the MinDef partition, each basis function χ_i is considered to be centered on a nucleus of the molecule. If, besides this, χ_i is a Spherical or Cartesian Gaussian, $\hat{\omega}_k^{ab}$ in this partition is defined by

$$\hat{\omega}_k^{ab} = [m_k(a)\Theta(\zeta_a - \zeta_b) + m_k(b)\Theta(\zeta_b - \zeta_a)], \tag{32}$$

where $m_k(a) = 1(0)$ if χ_a is (is not) centered at the nucleus of atom k, ζ_a and ζ_b are the exponents of χ_a and χ_b, respectively, and $\Theta(x)$ is the Heaviside step function $\left(\Theta(x>0)=1, \Theta(x<0)=0, \Theta(x=0)=\frac{1}{2}\right)$. The $\hat{\omega}_k^{ab}$ values follow from the following considerations. A product $\chi_a\chi_b$, with χ_a and χ_b centered at the nucleus of the same atom (say A), is assigned to this atom, so that $\hat{\omega}_k^{ab} = 1$ or 0 for $k = A$ and $k \neq A$, respectively. If χ_a is centered at the nucleus of atom A and χ_b is centered at the nucleus of a different atom B, $\chi_a\chi_b$ can always be written as a linear combination of Gaussians centered at a point $\boldsymbol{P}$ along the line joining these two nuclei. In this case, $\chi_a\chi_b$ is assigned to A or B depending on whether $\boldsymbol{P}$ is closer to the nucleus of A ($\zeta_a > \zeta_b$) or B ($\zeta_a < \zeta_b$), or half-and-half to A and B if $\boldsymbol{P}$ is just in the middle of the two nuclei ($\zeta_a = \zeta_b$). According to this, $\hat{\omega}_k^{ab} = 1(0)$ for $k = A$ $(k \neq A)$ if $\zeta_a > \zeta_b, \hat{\omega}_k^{ab} = 1(0)$ for $k = B$ $(k \neq B)$ if $\zeta_a < \zeta_b$, and $\hat{\omega}_k^{ab} = \frac{1}{2}$ for $k = A$ or $k = B$ if $\zeta_a = \zeta_b$.

Mulliken's partition is also given by Eq. 32 by choosing $\hat{\omega}_k^{ab} = 0, \frac{1}{2}, 1$ depending on whether none, one (either χ_a or χ_b) or both (χ_a and χ_b) primitive Gaussians are centered at the nucleus of atom k. In this partition, $(\mathbf{S}_{ij})_{\Omega_k}$ may also be written as

$$(\mathbf{S}_{ij})_{\Omega_k} = \sum_{a \in k} \sum_b C_{ai} C_{bj} \Sigma_{ab} = \sum_{a \in k} \left(\boldsymbol{P}^{ij}\Sigma\right)_{aa}, \tag{33}$$

where $\boldsymbol{P}^{ij} = C_{ai}C_{bj}$. However, in addition to the well-known shortcomings of Mulliken's population analysis, Mulliken's orbital space partition has the further disadvantage of giving non-symmetric AOM's, i.e. $(\mathbf{S}_{ij})_{\Omega_k} \neq (\mathbf{S}_{ji})_{\Omega_k}$, in general. The asymmetry of $\mathbf{S}_{\Omega_k}$ is solved within Löwdin's orbital space partition, where

$$\left(\mathbf{S}_{ij}\right)_{\Omega_k} = \sum_{a \in k} \left(\Sigma^{\frac{1}{2}} \boldsymbol{P}^{ij} \Sigma^{\frac{1}{2}} \right)_{aa}. \tag{34}$$

It is worth mentioning that Eqs. 33 and 34 are almost similar to the expressions used to obtain the net electronic charge of the Ω_k domain according to Mulliken and Löwdin population analyses, respectively. The only difference is that the $\boldsymbol{P}^{ij}$ elements must be substituted by $\boldsymbol{P}$, the full density matrix element associated with the natural MO of the systems, and defined by $\boldsymbol{P}_{ab} = \sum_i^{\mathrm{occ}} n_i C_{ai} C_{bi}$, where n_i is the electron population of the natural MO φ_i.

3 Chemical bonding concepts from the electron number distribution function

The chemical bonding concepts derived from the EDF computed for a QTAIM-based partition of the physical space may be readily generalized to deal with diffuse or orbital space partitions. For instance, the average number of electrons associated with a domain Ω_k is given by [7, 9]

$$\langle n_k \rangle = \sum_S n_k p(S) = \sum_S n_k p(n_1, \ldots, n_k, \ldots, n_m). \tag{35}$$

For an exhaustive partition, where w_k is the step function given in Eq. 3, $\langle n_k \rangle$ coincides exactly with the electron density en within the Ω_k region, i.e.

$$\sum_S n_k p(S) = \int_{\Omega_k} \rho(\boldsymbol{r}) d\boldsymbol{r}. \tag{36}$$

For diffuse partitions, where the boundaries of the Ω_k regions are not sharply defined, $\langle n_k \rangle$ is equal to the integral over $\mathcal{R}^3$ of an electron density associated with Ω_k, and defined by $\rho_k(\boldsymbol{r}) = w_k(\boldsymbol{r})\rho(\boldsymbol{r})$, i.e.

$$\sum_S n_k p(S) = \int \rho_k(\boldsymbol{r}) d\boldsymbol{r}. \tag{37}$$

Similarly, $\langle f(S) \rangle$, the mean value of an arbitrary function of the n_i's that define the resonance structure S is obtained as the statistical average:

$$\langle f(S) \rangle = \sum_S f(S) p(S). \tag{38}$$

Particularly relevant $f(S)$ functions are the multi-center delocalization indices [19–22], defined here as

$$\delta^{\Omega_a,\Omega_b,\ldots,\Omega_c} = \frac{(-2)^{K-1}}{(K-1)!} \times \langle (n_a - \langle n_a \rangle) \times (n_b - \langle n_b \rangle) \cdots (n_c - \langle n_c \rangle) \rangle, \tag{39}$$

where K is the number of domains involved in $\delta^{\Omega_a,\Omega_b,\ldots,\Omega_c}$. $\delta^{\Omega_a,\Omega_b} = -2(n_a - \langle n_a \rangle)(n_b - \langle n_b \rangle)$ corresponds to the two-fragments delocalization index, a measure of the number of pairs of electrons shared by Ω_a and Ω_b [12–15, 47–49]. Usually, in exhaustive partitions of $\mathcal{R}^3$, $\delta^{\Omega_a,\Omega_b}$ is computed from the expression

$$\delta^{\Omega_a,\Omega_b} = -2 \int_{\Omega_a} \int_{\Omega_b} [\rho_2(\boldsymbol{r}_1, \boldsymbol{r}_2) - \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)] d\boldsymbol{r}_1 d\boldsymbol{r}_2, \tag{40}$$

where $\rho_2(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is the second-order reduced density matrix (2-RDM). However, for multi-determinant wave functions, the 2-RDM is not generally available from standard computational packages. For this reason, Eq. 38, valid for any number of domains, real or orbital space partitions, and single- as well as multi-determinant wave functions, is a very valuable expression in the analysis of the chemical bond.

4 Computational details and test systems studied

We have computed the EDF of several molecules using both the exhaustive QTAIM space partition and the fuzzy as well as orbital-based partitions defined in Sect. 2. The systems range from covalent molecules (H_2, N_2) to others that are usually considered as ionic (NaCl). EDF for the family of the saturated second-row hydrides AH_n (A=Li,Be,B,C,N,O,F) have also been computed in order to explore how they evolve with the degree of polarization of the A–H bond. All the molecular wave functions of this work have been obtained with the `GAMESS` code [50]. Both Hartree-Fock (HF) and complete active space SCF (CASSCF) wave functions (CAS[i, j], i active electrons and j active orbitals) have been obtained for all the molecules except NaCl, where only the HF wave function was obtained. The CASSCF descriptions used are CAS[2,2] for H_2, CAS[10,8] for N_2, CAS[4,6] for LiH, CAS[6,7] for BeH_2, CAS[8,8] for BH_3, and CAS[10,9] for CH_4, NH_3, H_2O, and HF. A full Configuration Interaction (full-CI) wave function was also generated for H_2 for R(H-H) distances from 0.6 to 10.0 a_0. The basis sets used where Pople's 6-311G(d,p), except for N_2, in which a TZV (d,p) basis set was employed. Numerical radial and angular quadratures previously described [7] have been used to compute the atomic overlap matrices corresponding to the QTAIM and fuzzy partitions by using our `PROMOLDEN` code [51]. The AOM's of the orbital-based partitions (Mulliken, Löwdin, and MinDef) were analytically computed by using the McMurchie and Davidson algorithm [52]. Finally, all EDF computations have been performed by using our `EDF` code [9].

5 Results and discussion

5.1 H_2 molecule

The wave function of the H_2 molecule (H_a–H_b) at the HF level is a single Slater determinant made of the spin-orbitals

$\varphi_1\alpha$ and $\varphi_1\beta$ with $\varphi_1 = 1\sigma_g$. The generalized AOM is blocked into two similar (1×1) blocks. Regardless of the space partition and basis set used in the calculation, one has $(\mathbf{S}^\sigma_{11})_{\Omega_a} = (\mathbf{S}^\sigma_{11})_{\Omega_b} = \frac{1}{2}, p_\sigma(0,1) = p_\sigma(1,0) = \frac{1}{2} (\sigma = \alpha, \beta)$, $p(0,2) = p(2,0) = \frac{1}{4}$, and $p(1,1) = \frac{1}{2}$ in all of the cases. More interesting are the EDF corresponding to the CAS[2,2] calculation. Now, the wave function is a linear combination of two Slater determinants, $\Psi = d_1|1\sigma_g\alpha 1\sigma_g\beta| + d_2|1\sigma_u\alpha 1\sigma_u\beta|$, and the AOM in Ω_a within each spin block is

$$(\mathbf{S}^\sigma)_{\Omega_a} = \begin{pmatrix} \frac{1}{2} & a_{gu} \\ a_{ug} & \frac{1}{2} \end{pmatrix}, \qquad (41)$$

where $a_{gu} = \langle 1\sigma_g|1\sigma_u\rangle_{\Omega_a}$ and $a_{ug} = \langle 1\sigma_u|1\sigma_g\rangle_{\Omega_a}$. $(\mathbf{S}^\sigma)_{\Omega_b}$ is equal to $(\mathbf{S}^\sigma)_{\Omega_a}$ changing a_{gu} and a_{ug} by their respective negative values.

In all the partitions except in Mulliken's one, we have $a_{gu} = a_{ug}$, so that a single AOM element controls, thus, the differences between the EDF. It can be shown from Eqs. 14, 17, 27, and 28 that $p(0, 2) = p(0, 1;0, 1) = p(1, 0;1, 0) = p(2, 0)$, the probability of having simultaneously the two electrons in the same domain (Ω_a or Ω_b) is given by $\frac{1}{4} + 2d_1d_2a_{gu}^2$ ($\frac{1}{4} + d_1d_2a_{gu}^2 + d_1d_2a_{ug}^2$ in Mulliken's partition). The terms involving d_1 and d_2 are negative since these two coefficients have opposite signs, so that the greater the orthogonality between MO $1\sigma_g$ and $1\sigma_u$ in Ω_a, the greater the probability of finding both electrons in the same domain. Our results for a_{gu}, $p(0, 2) = p(2, 0)$, $p(1, 1)$, and $\delta^{\Omega_a,\Omega_b}$ using the six space partitions discussed in Sect. 2 are collected in Table 1. At the HF level, $p(2, 0)$ is simply $p_\alpha(1, 0) \times p_\beta(1, 0)$, i.e. both (opposite spin) electrons are not correlated. In other words, the probability that the β electron is in Ω_a when it is known for sure that the α electron is already in Ω_a is exactly 0.5. From Table 1, we observe that Hirshfeld partition tends to uncorrelate both electrons, the contrary happens in the MinDef partition, and Becke and Löwdin EDF are very similar to QTAIM EDF. The last column in the Table reflects the well-known fact that a small correlation between both electrons tends to overestimate the inter-fragment delocalization index $\delta^{\Omega_a,\Omega_b}$, approaching its value to that given by the HF model (1.0). Finally, as a consequence of the close to unit value of a_{ug}, the two opposite spin electrons are very correlated in Mulliken's partition, with relatively high and small values of $p(1, 1)$ and $\delta^{\Omega_a,\Omega_b}$, respectively.

Table 1 CAS[2,2] EDF for H_2 using different space partitions

EDF	a_{gu}	$p(0, 2) = p(2, 0)$	$p(1, 1)$	$\delta^{\Omega_a,\Omega_b}$
QTAIM	0.4373	0.2083	0.5833	0.8332
Becke	0.4145	0.2126	0.5749	0.8502
Hirshfeld	0.3037	0.2299	0.5402	0.9196
Mulliken	0.2051	0.1365	0.7270	0.5460
Löwdin	0.4030	0.2255	0.5490	0.9019
MinDef	0.6562	0.1561	0.6877	0.6245

$p(0, 2) = p(0, 1;0, 1)$ is the probability that 0 and 2 electrons are in the left (Ω_a) and right (Ω_b) hydrogens, respectively, and $p(1, 1) = p(1, 0;0, 1) + p(0, 1;1, 0)$, where $p(1, 0;0, 1)$ is the probability of having one α electron in Ω_a and one β electron in Ω_b, and $p(0, 1;1, 0)$ the probability of having one β electron in Ω_a and one α electron in Ω_b. The AOM element a_{gu} is defined as $\langle 1\sigma_g|1\sigma_u\rangle_{\Omega_a}$, and $a_{ug} = \langle 1\sigma_u|1\sigma_g\rangle_{\Omega_a}$ is equal to a_{gu} in all except Mulliken's partition, where $a_{ug} = 0.9998$

5.2 N_2 molecule

The HF and CAS[10,8] EDF for the N_2 (N_a–N_b) molecule are plotted in Fig. 1. $p(n_a \leq 3,\ n_b \geq 11)$'s have been excluded from it, as they are very small in all partitions. Similarly, since $p(n_a, n_b) = p(n_b, n_a)$, only EDF with $n_a \leq n_b$ are shown in the figure. As a first result, we notice that HF EDF are nearly independent of the partition used in their calculation, whereas this is not true for correlated EDF. Interestingly, the correlated $p(6,8)$ value, corresponding to the ionic N^+–N^- structure, is almost the same in all partitions. This $p(6,8)$ value marks a crossing point in the figure, as the correlated $p(7,7)$ value (corresponding to the neutral N^0–N^0 structure) increases in the order Löwdin $\simeq$ Hirshfeld < Becke < QTAIM < Mulliken < MinDef, whereas the correlated $p(5,9)$ value (N^{2+}-N^{2-} structure) decreases in the same order. Moreover, correlated $p(6,8)$'s are only marginally greater than their respective HF values, correlated $p(7,7)$'s are clearly greater than the HF $p(7,7)$'s, and correlated $p(5,9)$'s are marginally (Löwdin, Hirshfeld)

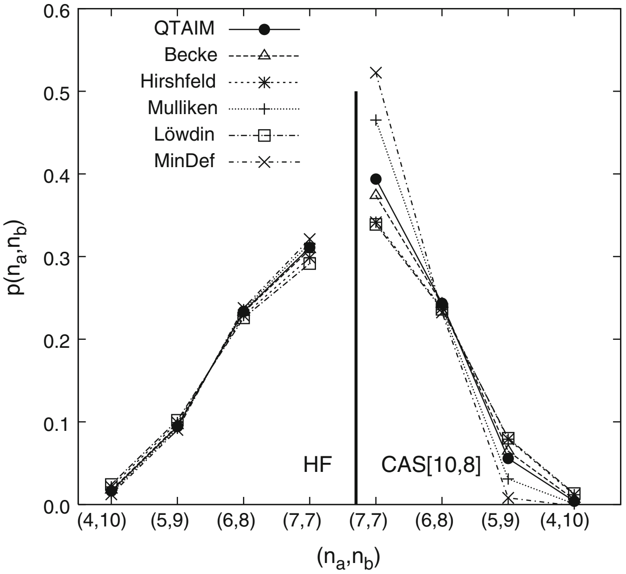

Fig. 1 Hartree-Fock and correlated EDF for N_2 according to different space partitions

Table 2 Hartree-Fock and correlated $p(7,7)$ values and delocalization indices, $\delta^{\Omega_a,\Omega_b}$, for the N_2 molecule

	HF		CAS[10,8]	
	$p(7,7)$	$\delta^{\Omega_a,\Omega_b}$	$p(7,7)$	$\delta^{\Omega_a,\Omega_b}$
QTAIM	0.3109	3.0408	0.3937	2.0113
Becke	0.3083	3.1073	0.3738	2.2298
Hirshfeld	0.2987	3.3664	0.3415	2.6879
Mulliken	0.3160	2.9110	0.4651	1.4618
Löwdin	0.2915	3.5613	0.3389	2.7703
MinDef	0.3211	2.7834	0.5225	0.9998

or clearly smaller (Becke, QTAIM, Mulliken, MinDef) than HF $p(5,9)$'s. This behavior points out that electron correlation tends to enhance the probability of neutral structures with respect to the HF calculation, at the expense of a decrease in the probabilities of the ionic ones. This fact is of general validity, at least in the QTAIM partition [7], the only one for which EDF have been computed up to now. As a direct consequence, electron delocalization is less pronounced when electron correlation is taken into account. This gives a correlated EDF, which is narrower than the corresponding HF EDF, and, consequently, smaller values for the covalency index, $\delta^{\Omega_a,\Omega_b}$, as it can be seen in Table 2. From this table, it is also evident that $\delta^{\Omega_a,\Omega_b}$ increases when $p(7,7)$ decreases and *vice versa*, both in the HF and correlated calculations. Besides this, and according to the above discussion, the HF $p(7,7)$ and $\delta^{\Omega_a,\Omega_b}$ values do not differ too much in the different methods, whereas they span a wide range of values in the CAS[10,8] calculation.

5.3 AH_n (A=Li,Be,B,C,N,O,F) hydrides

In this series of molecules, it seems reasonable to expect that the total charge of the central atom A, given by $q_A = Z_A - \langle n_A \rangle$, where Z_A is the nuclear charge and $\langle n_A \rangle$ the average number of electrons of this atom, as given by Eq. 35, starts being positive for LiH and ends being negative for FH. However, the actual values of q_A or $\langle n_A \rangle$, being directly dependent on the EDF of the molecule through Eq. 35 might be very different depending on the partition used in the calculation. For brevity, we will only discuss in this article the results for LiH, FH, and H_2O, and in the last case, we will use a partition into three domains, one for each of the atoms in the molecule. When only a two-domain partition is performed, with $\Omega_1 = A$ and $\Omega_2 = H_1 + \cdots + H_n$, the results and conclusions derived for LiH and FH may serve to (approximately) interpolate the results and conclusions for the other hydrides.

Table 3 EDF for LiH according to different space partitions

n_{Li}	QTAIM	Becke	Hirshfeld	Mulliken	Löwdin	MinDef
1	0.0065	0.0106	0.0083	−0.0002	0.0005	0.0001
	–	–	–	–	–	–
	0.0064	0.0101	0.0078	−0.0001	0.0004	0.0001
2	0.8948	0.7427	0.4860	0.4630	0.3188	0.7052
	0.9007	0.7500	0.4871	0.4744	0.3185	0.7052
	0.8905	0.7117	0.4477	0.3463	0.2760	0.6187
3	0.0961	0.2290	0.4160	0.4350	0.4913	0.2690
	0.0968	0.2320	0.4216	0.4288	0.4917	0.2691
	0.1012	0.2639	0.4691	0.5932	0.5500	0.3713
4	0.0026	0.0177	0.0897	0.1022	0.1895	0.0257
	0.0026	0.0179	0.0912	0.0969	0.1898	0.0257
	0.0018	0.0143	0.0754	0.0607	0.1737	0.0099
q_{Li}	0.9053	0.7462	0.4130	0.3605	0.1303	0.6798
	0.8981	0.7321	0.3959	0.3775	0.1288	0.6796
	0.9015	0.7176	0.3881	0.2853	0.1031	0.6091
$\delta^{Li,H}$	0.2079	0.4922	0.8772	0.8691	0.9864	0.5385
	0.1935	0.4641	0.8433	0.8575	0.9834	0.5382
	0.2108	0.5032	0.8081	0.6501	0.8811	0.5161

For each n_{Li} and partition, the first entry is the exact HF result, the second one the approximate probability that results after performing an isopycnic localization of HF MO and excluding the 1s Li core electrons from the calculation, and the third one the exact probability from a CAS[4,6] calculation

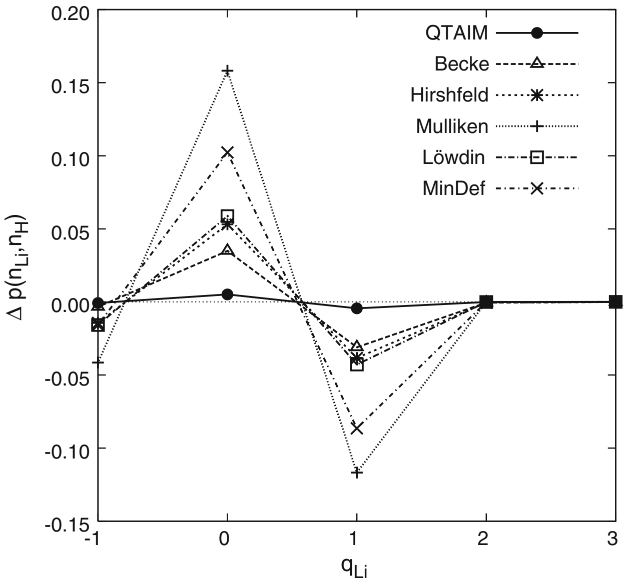

Fig. 2 EDF change of LiH in passing from the HF to the correlated CAS[4,6] calculation

Our results for LiH are collected in Table 3, and the EDF change in going from the HF to the correlated calculation is plotted in Fig. 2. As anticipated, the q_{Li} values in Table 3 denote that LiH is almost exclusively represented by the ionic $\mathrm{Li^+H^-}$ structure ($p(n_{\mathrm{Li}} = 2, n_{\mathrm{H}} = 2) = 0.895$, HF) in the QTAIM partition, but the $\mathrm{Li^+H^-}$ and the neutral $\mathrm{Li^0H^0}$ structures have significant probabilities in all the other partitions. In this sense, Becke's fuzzy partition is the one giving the closest results to QTAIM. The highly ionized $\mathrm{Li^{2+}H^{2-}}$ structure has a negligible probability in all the cases. The negative value of $p(1{,}3)$ in the HF Mulliken partition is due to the non-symmetric character of the AOM in this method. It should be stressed that $p(1{,}3)$, corresponding to the unrealistic $\mathrm{Li^-H^+}$ structure, has an appreciable value in all but QTAIM partition. This fact, together with the comparable importance of the $\mathrm{Li^+H^-}$ and $\mathrm{Li^0H^0}$ structures, justifies the low values found for q_{Li} in these cases.

The second entry in Table 3 illustrates how EDF depend on the use of the core approximation. An algebraic analysis of the equations given in Sect. 2 shows that if $(\mathbf{S}_{ii})_{\Omega_k} = \langle\varphi_i|\varphi_i\rangle_{\Omega_k} \simeq 1$, i.e. a diagonal element of the AOM is close to 1.0 within a domain Ω_k, then the φ_i orbital can be safely excluded from the EDF computation. The closest to 1.0 is $(\mathbf{S}_{ii})_{\Omega_k}$ the better the approximation of ignoring φ_i is. One can say that φ_i is a core orbital of domain Ω_k. To analyze numerically how this approximation works in LiH, we have transformed its two HF MO by means of an isopycnic (unitary) localization that does not modify the wave function. One of the rotated or isopycnic MO (say φ_{loc}) has $\langle\varphi_{\mathrm{loc}} \mid \varphi_{\mathrm{loc}}\rangle_{\mathrm{Li}} > 0.99$ in all partitions. Hence, φ_{loc} is almost fully localized on the Li atom and can thus be identified with the 1s Li core. Ignoring φ_{loc} transforms LiH into a molecule with only two effective electrons (one α electron and one β electron), both placed on the second isopycnic MO (say φ_{del}). Since $p(2,\,0) = p_\alpha(1,\,0) \times p_\beta(1,\,0)$, $p(0,\,2) = p_\alpha(0,\,1) \times p_\beta(0,\,1)$, and $p(1,\,1) = p_\alpha(0,\,1) \times p_\beta(1,\,0) + p_\alpha(1,\,0) \times p_\beta(0,\,1)$ at the HF level, the EDF in LiH is determined by a single AOM element $n_d = \langle\varphi_{\mathrm{del}} \mid \varphi_{\mathrm{del}}\rangle_{\mathrm{Li}}$. Actually, in the QTAIM partition, φ_{del} is very localized on the H atom since $n_d = 0.0510$, so that $p(2,0) = n_d^2 = 0.0026$, $p(0,2) = (1 - n_d)^2 = 0.9006$, and $p(1{,}1) = 2n_d(1 - n_d) = 0.0968$, i.e. both electrons of φ_{del} are with a probability greater than 90% in the H domain. The above three numbers, also collected in Table 3, are in very good agreement with the exact values, obtained without using the core approximation. In the other five partitions, φ_{del} is not so localized on the H atom so that $p(2{,}0)$ and $p(0{,}2)$ are non-negligible. However, the core approximation works as well as in the QTAIM partition. Assuming that $\langle\varphi_{\mathrm{loc}}|\varphi_{\mathrm{loc}}\rangle_{\mathrm{Li}}$ is exactly 1.0, it can be shown that $\delta^{\mathrm{Li,H}} = 4n_d(1 - n_d)$. The $\delta^{\mathrm{Li,H}}$ values obtained using this formula (second entry in Table 3) are in fact very close to the exact HF results, showing again that only two electrons partitipace in the Li–H chemical bond. The small differences are due to the residual contribution of the extremely localized MO, φ_{loc}, to $\delta^{\mathrm{Li,H}}$.

The third entry in Table 3 corresponds to the correlated CAS[4,6] calculation and the $\Delta p(n_{\mathrm{Li}}, n_{\mathrm{H}}) = p_{\mathrm{CAS}} - p_{\mathrm{HF}}$ differences are shown in Fig. 2. Correlation effects are very small in the QTAIM partition, moderate in Becke's partition, and progressively more important in the order Hirshfeld < Löwdin < MinDef < Mulliken. There seems that orbital space partitions exaggerate the importance of electron correlation on the computation of the EDF. In all the cases, the probability of the neutral (ionic) structure $\mathrm{Li^0H^0}$ ($\mathrm{Li^+H^-}$) increases (decreases) when electron correlation is taken into account. As a consequence, the HF n_{Li} (q_{Li}) values are greater (smaller) than the CAS[4,6] values, and "correlated' LiH becomes less ionic than "HF"LiH. The smaller delocalization of the valence 2s Li electron on the H domain produces a noticeable decrease in $\delta^{\mathrm{Li,H}}$ in going from the HF to the correlated calculation in all except the QTAIM and Becke's partitions. In fact, in these two cases, the CAS[4,6] $\delta^{\mathrm{Li,H}}$ value is slightly greater than the HF one.

The results for FH are given in Table 4. In QTAIM, the $\mathrm{F^-H^+}$ structure is almost four times more probable ($p(10{,}0) = 0.77$) than $\mathrm{F^0H^0}$ ($p(9{,}1) = 0.22$). As in LiH, the ionic structure is not so dominant over the neutral one in the rest of partitions. Actually, the structure $\mathrm{F^+H^-}$, with one of the electrons of F atom transferred to H atom, has a non-negligible probability in all except the QTAIM method. According to the EDF values, the total charge of F atom, q_{F}, goes from $q_{\mathrm{F}} = -0.749$ in QTAIM to almost 0.0 in Löwdin ($q_{\mathrm{F}} = -0.099$). Overall, the electron delocalization in fuzzy and orbital space partition methods is much higher than in the QTAIM method. As a result, $\delta^{\mathrm{F,H}}$ is much higher in the former, with unrealistic values of 1.259 and 1.266 in the Hirshfeld and Löwdin methods, respectively. Including electron correlation effects increases (decreases) a little the probability of the neutral (ionic

Table 4 HF (first entry) EDF for FH according to different space partitions

n_{F}	QTAIM	Becke	Hirshfeld	Mulliken	Löwdin	MinDef
8	0.0168	0.0882	0.1411	0.1163	0.1845	0.0956
Δ	−0.0003	−0.0099	−0.0005	−0.0199	−0.0073	−0.0302
9	0.2168	0.3899	0.4065	0.4371	0.4481	0.4320
Δ	0.0330	0.0273	0.0232	0.0562	0.0260	0.0755
10	0.7663	0.5160	0.4274	0.4425	0.3402	0.4741
Δ	−0.0326	−0.0155	−0.0187	−0.0342	−0.0134	−0.0452
n_{F}	9.7491	9.4159	9.2343	9.3179	9.0994	9.3821
Δ	−0.0323	−0.0015	−0.0094	−0.0103	0.0053	−0.0149
$\delta^{\mathrm{F,H}}$	0.4454	0.9110	1.2494	0.9492	1.2663	0.8318
Δ	0.0285	−0.0648	−0.0699	−0.1122	−0.0936	−0.1281

Rows labeled Δ stand for the changes in passing from the HF to the correlated CAS[10,9] results

Table 5 HF (first entry) and CAS[10,9] (second entry) EDF results for the H_2O molecule according to different space partitions

n_O	n_{H_1}	n_{H_2}	QTAIM	Becke	Hirshfeld	Mulliken	Löwdin	MinDef
8	1	1	0.0970	0.1784	0.1797	0.2130	0.2122	0.2310
			0.1320	0.2113	0.2011	0.3168	0.2436	0.3407
9	1	0	0.2033	0.1937	0.1638	0.1829	0.1482	0.1730
			0.2249	0.2048	0.1661	0.1930	0.1504	0.1820
7	1	2	0.0125	0.0476	0.0680	0.0643	0.0895	0.0769
			0.0131	0.0430	0.0694	0.0518	0.0895	0.0614
8	0	2	0.0273	0.0546	0.0687	0.0560	0.0659	0.0579
			0.0276	0.0504	0.0682	0.0455	0.0628	0.0456
10	0	0	0.4134	0.2020	0.1385	0.1554	0.1000	0.1291
			0.3339	0.1734	0.1197	0.0955	0.0847	0.0758
q(O)			−1.2029	−0.6482	−0.3232	−0.4996	−0.1646	−0.4022
			−1.0873	−0.6384	−0.3064	−0.4608	−0.1677	−0.3807
δ^{O,H_1}			0.6479	0.9634	1.2079	0.9625	1.1637	0.9502
			0.6367	0.8407	1.1064	0.6832	1.0400	0.6703

F^-H^+) structure in all the cases, predicting a lesser H to F electron transfer and (except in QTAIM), a smaller $\delta^{F,H}$ value.

The EDF results of FH when the core approximation is used are not included in Table 4. However, it is interesting to comment that, despite the total number of electrons in FH is much higher than in LiH, the isopycnic localization of its five canonical HF MO leaves, in the case of the QTAIM partition, an unique isopycnic valence MO ($\varphi_{\rm del}$), the other four isopycnic MO $\left(\varphi^i_{\rm loc}, i = 1-4\right)$ being almost completely localized on the F atom, with $\langle \varphi^i_{\rm loc}|\varphi^i_{\rm loc}\rangle_F$ values greater than 0.99. This means that the QTAIM results in Table 4 may be explained in terms of only two (α and β) valence electrons, both placed in the $\varphi_{\rm del}$ MO. The overlap integral $\langle \varphi_{\rm del}|\varphi_{\rm del}\rangle_F$ takes a value of 0.8818, which (using the same arguments as in the LiH case) leads to $p(8, 2) = 0.0141$, $p(9, 1) = 0.2091$, and $p(10,0) = 0.7768$, in fairly good agreement with the *all-electron* results of Table 4. The core approximation has also been used with the fuzzy and orbital-based partitions, and the resulting EDF also agree relatively well with the *all-electron* results. However, the higher electron delocalization found in these methods prevents the isopycnic localization to be as effective as in the QTAIM partition: Only two (Becke, Mulliken, MinDef) or one (Hirshfeld, Löwdin) isopycnic MO have $\langle \varphi^i_{\rm loc}|\varphi^i_{\rm loc}\rangle_F$ values greater than 0.99 and can be considered as F core orbitals.

The computed EDF for H_2O using three-domain partitions are collected in Table 5. The ionic structure $H_1^+H_2^+O^{2-}$, corresponding to a full transfer of the H_1 and H_2 electrons to the O atom, is the most probable one in the QTAIM partition, followed by $H_1^0H_2^+O^- \equiv H_1^+H_2^0O^-$, and $H_1^0H_2^0O^0$. In the other partitions, the structure $H_1^0H_2^0O^0$ (Hirshfeld, Mulliken, Löwdin, MinDef) or $H_1^0H_2^+O^- \equiv H_1^+H_2^0O^-$ (Becke) dominates over the neutral atoms one. A feature of fuzzy-atom and orbital-based partitions, also found in other molecules, is that the number of resonance structures having a significant probability is greater than in the QTAIM method. In the present case, this fact is clearly illustrated in Fig. 3, where we have plotted the HF values of the 10 most probable structures (not necessarily the same in all the methods). The EDF is considerably narrower when the QTAIM partition is used. When correlation effects are included, the QTAIM EDF widens slightly, making more probable the $H_1^0H_2^0O^0$ and $H_1^0H_2^+O^- \equiv H_1^+H_2^0O^-$ structures, and less the $H_1^+H_2^+O^{2-}$ one. The same happens in the other methods, so that the total charge of the O atom becomes less negative (except in Löwdin method) when electron correlation is included in the calculation.

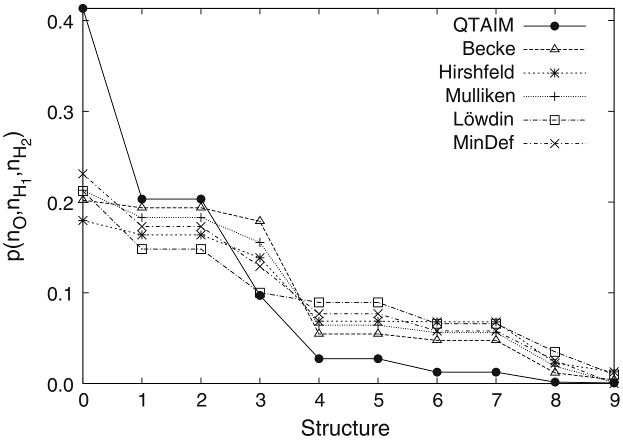

Fig. 3 $p(n_O, n_{H_1}, n_{H_2})$ values for the 10 most probable structures of H_2O molecule according to different space partitions

Concerning the delocalization index $\delta^{\mathrm{O,H_1}}$, we observe in Table 5 that, as a consequence of the too wide Hirshfeld and Löwdin EDF, these two methods predict an unrealistic (greater than 1.0) value for this chemical bonding indicator. Becke, Mulliken, and MinDef partitions give a very similar $\delta^{\mathrm{O,H_1}}$ (around 0.95–0.96). Although smaller than 1.0, this value is also probably too high for this (traditionally considered) polar O–H bond. In the HF QTAIM partition, $\delta^{\mathrm{O,H_1}} \simeq 0.65$, this value decreasing a little in the correlated calculation due to the higher probability of the neutral atoms structure, i.e. to the weakening of electron delocalization of the H_1 and H_2 electrons onto the O atom. This effect is even stronger in the other partitions.

5.4 NaCl molecule

Our HF EDF results for NaCl, a classical ionic molecule where a large fraction of the valence 3s Na electron is transferred to the Cl atom, are plotted in Fig. 4. The ionic structure, $\mathrm{Na^+Cl^-}$, dominates over the neutral atoms one, $\mathrm{Na^0Cl^0}$, in all the cases. However, as in all previous examples, there are clear differences between the different methods. Löwdin's partition predicts the least ionic NaCl molecules followed by Hirshfeld, Mulliken, Becke, MinDef, and QTAIM methods. Contrary to chemical intuition, some diffuse and/or orbital-based methods (particularly Löwdin's) predict a non- negligible probability for the $\mathrm{Na^-Cl^+}$ structure. The EDF is normalized, so that the large the probabilities of the $\mathrm{Na^-Cl^+}$, $\mathrm{Na^{2+}Cl^{2-}}$, and (above all) $\mathrm{Na^0Cl^0}$ structures are, the smaller the probability of the classical ionic structure $\mathrm{Na^+Cl^-}$ is. As a result, the Löwdin and Hirshfeld (and to a lesser extent Mulliken and Becke) methods give wider EDF than the QTAIM partition. The standard Hirshfeld's method, where the atomic densities of the neutral atoms Na and Cl are used to obtain the promolecular density, $\rho^0(\boldsymbol{r})$, gives a too high $p(\mathrm{Na^0,Cl^0})$ value, predicting a NaCl molecule much less ionic than expected. However, a Hirshfeld's EDF with the promolecular density given by $\rho^0(\boldsymbol{r}) = \rho_{\mathrm{Na^+}}(\boldsymbol{r}) + \rho_{\mathrm{F^-}}(\boldsymbol{r})$ is almost indistinguishable from that of the QTAIM partition. This fact evidences the well-known tendency of Hirshfeld's partition to give *atoms in molecules* atomic charges (q_{A}) fairly close to the q_{A}^0's, the atomic charges associated with the *in vacuo* atomic densities used to compute $\rho^0(\boldsymbol{r})$. In this sense, it is worth mentioning that an iterative Hirshfeld population analysis method has recently been proposed in which $q_{\mathrm{A}} = q_{\mathrm{A}}^0$ in convergence. The results obtained for a large set of molecules using this iterative Hirshfeld scheme are generally very close to the ones derived from the QTAIM.

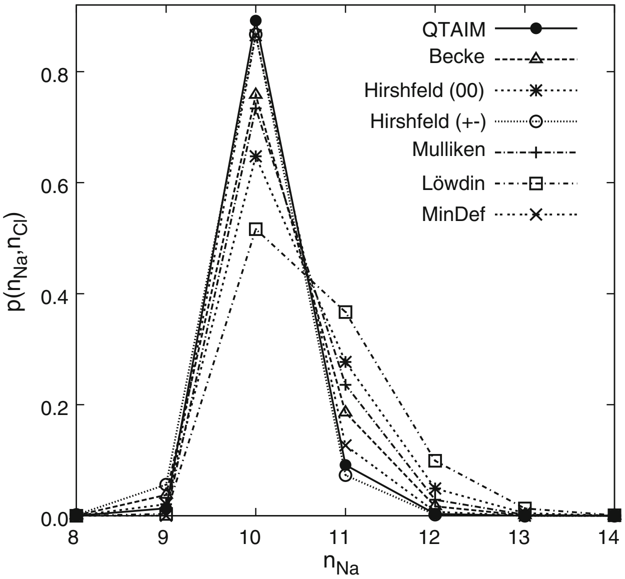

Fig. 4 HF EDF for the NaCl molecule according to different space partitions. (00) and (+-) in Hirshfeld partition mean that the reference atomic densities of (Na,Cl) neutral atoms and ($\mathrm{Na^+}$, $\mathrm{Cl^-}$) ions are used, respectively

6 Summary and conclusions

In this paper, the algebraic method previously derived to compute the electron number distribution functions for exhaustive partitions of the physical space into sharp-boundary atoms has been extended to allow for a space partition in terms of fuzzy atoms, as well as orbital-based partitions. The QTAIM, Becke, Hirshfeld, Mulliken, Löwdin, and MinDef methods, all of them described in the text, have been explicitly considered. They differ in the way the atomic overlap matrix elements, the main quantities from which the EDF depend, are computed from the molecular wave function. Once this matrix is available, all of them are included into a unified approach that only needs taking care of the non-symmetric character of this matrix in the particular case of Mulliken's partition.

The developed scheme has been applied to compare the EDF found using the above six partition methods for a series of covalent and ionic systems. The EDF for H_2, N_2, AH_n hydrides (A=Li,Be,B,C,N,O,F), and NaCl molecules have computed at the HF and (except in NaCl) CAS levels. Overall, our results show that EDF are significantly wider in fuzzy- and orbital-based methods than in the QTAIM partition. As a result, the atom–atom delocalization indices $\delta^{A,B}$ are usually higher in the former. In systems with an important degree of inter-atomic electron charge transfer, such as LiH, FH, and NaCl, fuzzy- and orbital-based methods have a tendency to decrease this transfer when compared to the exhaustive partition provided by the QTAIM method. The Hirshfeld's method that uses *in vacuo* ionic (instead of neutral atoms) densities is an exception to the above rule, providing EDF in good agreement with QTAIM EDF. In general, electron correlation has a limited effect over QTAIM EDF, but EDF from fuzzy- and orbital-based partitions are considerably more affected.

Acknowledgments The authors acknowledge financial support from the Spanish MICINN, Project No. CTQ2009-08376, the European Union FEDER funds, the MALTA-Consolider program (CSD2007-00045), and the FICYT (IB09-019). EF and AMP devote this work to the memory of our friend, colleague, and co-author Miguel A. Blanco, who passed away during the preparation of this manuscript.

References

1. McWeeny R (1992) Methods of molecular quantum mechanics, Chapter 14, 2nd edn. Academic Press, London
2. Bader RFW (1990) Atoms in molecules. Oxford University Press, Oxford
3. Martín Pendás A, Francisco E, Blanco MA (2007) J Phys Chem A 111:1084
4. Martín Pendás A, Francisco E, Blanco MA (2007) Chem Phys Lett 437:287
5. Gatti C, Macchi P (eds) (2010) Modern charge-density analysis, 1st edn. Springer, New York
6. Cancès E, Keriven R, Lodier F, Savin A (2004) Theor Chem Acc 111:373
7. Francisco E, Martín Pendás A, Blanco MA (2007) J Chem Phys 126:094102
8. Martín Pendás A, Francisco E, Blanco MA (2007) J Chem Phys 127:144103
9. Francisco E, Martín Pendás A, Blanco MA (2008) Comput Phys Commun 178:621
10. Martín Pendás A, Francisco E, Blanco MA (2007) Phys Chem Chem Phys 9:1087
11. Francisco E, Martín Pendás A, Blanco MA (2009) J Chem Phys 131:124125
12. Ángyán JG, Loos M, Mayer I (1994) J Phys Chem 98:5244
13. Fradera X, Austen MA, Bader RFW (1999) J Phys Chem A 103:304
14. Fradera X, Poater J, Simon S, Duran M, Solà M (2002) Theor Chem Acc 108:214
15. Poater J, Duran M, Solà M, Silvi B (2005) Chem Rev 105:3911
16. Cioslowski J, Mixon ST (1991) J Am Chem Soc 113:4142
17. Kar T, Ángyán JG, Sannigrahi AB (2000) J Phys Chem A 104:9953
18. Matito E, Solà M, Salvador P, Duran M (2007) Faraday Discuss 135:325
19. Mundim KC, Giambiagi M, de Giambiagi MS (1994) J Phys Chem 98:6118
20. Sannigrahi AB, Kar T (1999) Chem Phys Lett 299:518
21. Sannigrahi AB, Kar T (2000) J Mol Struct (Theochem) 496:1
22. Bochicchio RC, Ponec R, Torre A, Lain L (2001) Theor Chem Acc 105:292
23. Martín Pendás A, Francisco E, Blanco MA (2007) Faraday Discuss 135:423
24. Martín Pendás A, Francisco E, Blanco MA, Gatti C (2007) Chem Eur J 13:9362
25. Tiana D, Francisco E, Blanco MA, Macchi P, Sironi A, Martín Pendás A (2010) J Chem Theory Comput 6:1064
26. Daudel R (1968) The fundamentals of theoretical chemistry. Pergamon Press, Oxford
27. Daudel R, Bader RFW, Stephens ME, Borrett DS (1974) Can J Chem 52:1310
28. Becke AD, Edgecombe KE (1990) J Chem Phys 92:5397
29. Silvi B, Savin A (1994) Nature 371:683
30. Martín Pendás A, Francisco E, Blanco MA (2008) Chem Phys Lett 454:396
31. Blanco MA, Martín Pendás A, Francisco E (2005) J Chem Theory Comput 1:1096
32. Martín Pendás A, Blanco MA, Francisco E (2007) J Comput Chem 28:161
33. Martín Pendás A, Blanco MA, Francisco E (2006) J Chem Phys 125:184112
34. Martín Pendás A, Blanco MA, Francisco E (2009) J Comput Chem 30:98
35. Francisco E, Martín Pendás A, Blanco MA (2006) J Chem Theory Comput 2:90
36. Francisco E, Martín Pendás A, Blanco MA, Costales A (2007) J Phys Chem A 111:12146
37. Mulliken RS (1955) J Chem Phys 23:1833
38. Löwdin PO (1950) J Chem Phys 18:365
39. Becke AD (1988) J Chem Phys 88:2547
40. Hirshfeld FL (1977) Theor Chim Acta 44:129
41. Ghosh DG, Biswas R (2002) Int J Mol Sci 3:87
42. Bultinck P, Van Alsenoy C, Ayers PW, Carbó-Dorca R (2007) J Chem Phys 126:144111
43. Kullback K, Leibler RA (1951) Ann Math Stat 22:79
44. Cioslowski J (1990) Int J Quantum Chem S24:15
45. Rico JF, López R, Ramírez G (1999) J Chem Phys 110:4213
46. Rico JF, López R, Ema I, Ramírez G (2002) J Chem Phys 117:533
47. Bader RFW, Stephens ME (1975) J Am Chem Soc 97:7391
48. Rincón L, Alvarellos JE, Almeida R (2005) J Chem Phys 122:214103
49. Rincón L, Alvarellos JE, Almeida R (2005) J Chem Phys 122:214104
50. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JA (1993) J Comput Chem 14:1347
51. Martín Pendás A (2010) (unpublished)
52. McMurchie LE, Davidson ER (1978) J Comput Phys 26:218