



A first step towards Quantum Energy Potentials of Electron Pairs

Julen Munárriz,^{a,b,*} Rubén Laplaza,^{a,b} A. Martín Pendás^c and Julia Contreras-García^{b,*}Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

A first step towards the construction of a quantum force field for electron pairs in direct space is taken. Making use of topological tools (Interacting Quantum Atoms and the Electron Localization Function), we have analyzed the dependency of electron pairs electrostatic, kinetic and exchange-correlation energies upon bond stretching. Simple correlations were found, and can be explained with elementary models such as the homogeneous electron gas. The resulting energy model is applicable to various bonding regimes: from homopolar to highly polarized and even to non-conventional bonds. Overall, this is a fresh approach for developing real space-based force fields including an exchange-correlation term. It provides the relative weight of each of the contributions, showing that, in common Lewis structures, the exchange correlation contribution between electron pairs is negligible. However, our results reveal that classical approximations progressively fail for delocalized electrons, including lone pairs. This theoretical framework justifies the success of the classic Bond Charge Model (BCM) approach in solid state systems and sets the basis of its limits. Finally, this approach opens the door towards the development of quantitative rigorous energy models based on the ELF topology.

Introduction

The simulation of big systems requires the use of analytical force fields (FFs) that enable the performance of molecular dynamic simulations in reasonable computing times. This is usually done by following simple atomic models. For example, bond lengths are commonly analysed in terms of the Hooke's law, where potential and kinetic energy balances are parametrized to adjust to experimental/computational results.^{1,2} Nonetheless, in many cases, chemical behaviour is determined by non-classical terms (*e.g.* exchange or correlation), which are not easily encoded in these potentials. We propose hereby a change of paradigm that is able to easily capture electron reorganization (towards building reactive FFs), as well as quantum terms.

Atomic models usually require specific parametrization for different coordination numbers. This stems from the fact that electrons (*i.e.* bonds) are the ones being re-allocated during a reaction. Therefore, the static view of electrons belonging to one atom, and hence, its oxidation state, can be avoided if the electron pair is taken as the main unit. It should be noted that this approach has been used for a long time in chemistry, although not in a quantitative manner. Before the advent of formal force fields, the use of mere electrostatic reasoning among electron pairs allowed for the first time to predict

molecular structures, in what is nowadays known as the Valence Shell Electron Pair Repulsion (VSEPR) theory.³

Moreover the use of electron pairs also provides a simpler link to quantum mechanics. The usefulness of electron pairs as the basis of the representation lies on the fact that paired electrons behave very much like bosons.^{4,5} This means that Pauli correlation is minimized in a system composed by perfectly localized electron pairs, so that the exchange-correlation term (more difficult to disentangle with simple models) is minimal. These advantages have been exploited in both, classical approaches, such as the Bond Charge Model (BCM),^{6,7} where only potential and kinetic energy terms are included; as well as in more updated formulations based on quantum mechanical calculations.⁸

In order to provide robust and rigorous analytical expressions for a quantum FF, it is necessary to count with a recipe for the identification of electron pairs, and their parametrization. For that, the Interacting Quantum Atoms (IQA) energy decomposition scheme is very suitable, since it allows to obtain and quantify the different energetic contributions from the real space partitions of the system.^{9,10} The advantages of the IQA scheme are showcased by its success in the study of different chemical phenomena, such as halogen and hydrogen bonding,^{11,12} and metal–ligand interactions,¹³ among many others. However, these calculations are expensive and limited to small systems. Therefore, by reconstructing models from this scheme, the advantages of IQA could be transferred to bigger systems.

For translating the IQA partition to electron pair energy models, an electron-pair real-space partition is needed. The IQA scheme is usually coupled to the Quantum Theory of Atoms in Molecules (QTAIM) partition in order to associate the different energy terms to atomic contributions. In this sense, it is worth

^a Departamento de Química Física and Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, 50009, Zaragoza, Spain

^b Sorbonne Université, CNRS, Laboratoire de Chimie Théorique, LCT, F. 75005 Paris, France.

^c Laboratorio de Química Física y Analítica, Universidad de Oviedo, 33006, Oviedo, Spain.

†Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

noting the previous work of Popelier *et al.*, who developed the first protein FF based on the QAIM topology.¹⁴ This framework also proved to be suitable for geometry optimizations when combined with machine learning approaches.^{15,16} However, since we want to focus on electron pairs, we will couple IQA to the Electron Localization Function (ELF) topology.¹⁷ The ELF is able to recover Lewis entities, as bonds and lone pairs, providing the classical VSEPR picture, and is thus the perfect starting point for our energy model, allowing to quantify (classic and quantum) interactions between electron pairs.¹⁸

In this article we propose a first step in constructing a stretching energy model based on the ELF topology. In order to do so, we evaluate electron pair interactions through classical (kinetic and electrostatic) and quantum energy terms (exchange and correlation). The ease and chemical rigor with which the model parameters can be obtained set the foundations of a fresh approach to real space force fields. In addition, a careful analysis of the relative weight of the classical and quantum energy terms is presented. Valuable implications in general force field development are also derived: we identify the chemical bonds that require quantum terms among those that can be studied under a purely classical perspective. The resulting model is beautifully connected with previous “intuitive” models (*e.g.* VSEPR, BCM). It also sets the basis for connecting the ELF topology with the energy of the system, leading to a simple chemical model that attempts to answer the initial question: “*how do electron pairs (bonds, lone pairs) contribute to the energy of the system?*”

Theoretical and Computational Methods

All energy scans and wavefunctions for ELF and IQA analyses were obtained through DFT calculations, using the Gaussian 09 program package.¹⁹ The B3LYP exchange-correlation functional^{20,21} was used, in conjunction with the 6-31G(*d*) basis set.^{22,23} Stability checks were carried out with other functionals (PBE and PBE0)^{24–26} and with Hartree-Fock, in conjunction with bigger basis sets (6-311G(*d,p*), cc-PVDZ and cc-PVTZ).²⁷ The results obtained by varying the computational conditions led to the same behaviour, yielding very similar fittings, while the computational cost increased considerably with the increase in the basis set size. These results are provided in the Supporting Information (see Tables S1–S6). Hence, the small basis set was favoured. Bonding descriptors were computed at the B3LYP/6-311G(*d,p*) level to further ensure basis set convergence of the density at the bond critical points.

The partition into localized electron pairs was carried out with the Electron Localization Function (ELF).¹⁷ The ELF kernel, χ_σ , can be interpreted as a measure of the excess of local kinetic energy due to the Pauli Principle, relative to the homogeneous electron gas kinetic energy density.²⁸ The ELF, $\eta(\vec{r})$, is mapped through a Lorentzian function (see eq. 1), to a scale ranging from 0 (when $\chi_\sigma \rightarrow \infty$), to 1 (when $\chi_\sigma \rightarrow 0$).

$$\eta(\vec{r}) = \frac{1}{1 + \chi_\sigma^2} \quad (1)$$

The gradient of this function, $\nabla\eta$, is used to induce a topological partition which divides the space into non-overlapping regions (basins) whose properties can be determined by integrating the appropriate densities over their associated volume. Hence, if one is interested in, for example, lone pairs populations, it suffices to integrate the electron density, ρ , over the corresponding region associated to the lone pair maximum.

The ELF studies were performed with a locally modified version of the TopMod program,²⁹ using the monodeterminantal B3LYP/6-31G(*d*) wavefunctions in conjunction with a cubic tridimensional grid of 200 points in each direction. ELF plots were made with Chimera software.^{30,31} The energies of the different topological basins can be calculated with the Interacting Quantum Atoms (IQA) energy decomposition scheme.³² This approach provides a set of unique and rigorous energetic terms that additively recover the exact energy of the system. Unlike many topological analyses, this method is not only suitable for stationary points (*e.g.* equilibrium geometries), such as with virial related energy partitions, but also for non-equilibrium geometries. This feature was crucial for evaluating the energy terms along bond elongations in geometry scans. The energy terms are calculated by partitioning the first and second-order density matrices with respect to real space partitions, usually the QAIM atomic basins (see eq. 2).

$$E = \frac{\sum_A (T_A + V_{ee}^{AA} + V_{en}^{AA})}{\sum_A E_{intra}^A} + \frac{\sum_{A>B} (V_{en}^{AB} + V_{ne}^{AB} + V_{nn}^{AB} + V_{ee}^{AB})}{\sum_{A>B} E_{inter}^{AB}} \quad (2)$$

This way, the total energy is divided into an intra-atomic contribution, E_{intra}^A , and the inter-atomic interaction between each pair of atoms, E_{inter}^{AB} . The former is the sum of the kinetic energy of electrons in atom A, T_A , and the electron-electron and electron-nucleus interactions, V_{ee}^{AA} and V_{en}^{AA} , respectively. E_{inter}^{AB} is the sum of the interaction among the electrons in A and the nucleus of B, V_{en}^{AB} ; the nucleus of A and the electrons in B, V_{ne}^{AB} ; and the nucleus-nucleus and electron-electron interactions between A and B, V_{nn}^{AB} and V_{ee}^{AB} , respectively.

The Electron-electron repulsion can be further divided into a classical (electrostatic) contribution, V_{elec}^{AB} , and a non-classical one, V_{XC}^{AB} , which is the sum of exchange (X) and correlation (Corr) terms (see eq. 3).

$$V_{ee}^{AB} = V_{elec}^{AB} + V_{XC}^{AB} = V_{elec}^{AB} + V_X^{AB} + V_{Corr}^{AB} \quad (3)$$

Following this scheme, each contribution to the total energy (intra and interatomic) can be expressed just like in atomistic approaches, that is, as a sum of kinetic and electrostatic terms, plus a non-classical interatomic term, as shown in eq. 4 and 5.

$$E_{inter}^{AB} = V_{Coul}^{AB} + V_{XC}^{AB} \quad (4)$$

$$E_{intra}^A = T_A + V_{Coul}^{AA} + V_{XC}^{AA} \quad (5)$$

Where all electrostatic terms have been put under a common “Coul” index: $V_{Coul}^{AB} = V_{nn}^{AB} + V_{elec}^{AB} + V_{en}^{AB} + V_{ne}^{AB}$ and $V_{Coul}^{AA} = V_{elec}^{AA} + V_{en}^{AA} + V_{ne}^{AA}$.

IQA can also be applied to ELF partitions. When considering bonding basins (and valence in general) within the IQA partition,

the nuclear terms presented in the previous equations become null ($V_{en}^{BB} = 0$, $V_{nn}^{AB} = 0$, $V_{en}^{AB} = 0$). Therefore, for a bonding basin, "Bond", which interacts with a core basin that represents an atom A, the different energy terms can be expressed as shown in eq. 6 and 7.

$$E_{intra}^{Bond} = T_{Bond} + V_{Coul}^{Bond} \quad (6)$$

$$E_{inter}^{A-Bond} = V_{Coul}^{A-Bond} + V_{XC}^{A-Bond} \quad (7)$$

Since we are interested in developing an energy model accounting for interactions between electron pairs, this was the selected approach. This way, it is possible to calculate interactions between atoms and "bonds", as well as interactions between two "bonds" and "bonds" with "lone pairs". For this purpose, the original code was modified so as to perform integration tasks over ELF basins.¹⁸ The IQA-ELF approach provides an accurate reference that can be used to analyse the behaviour of the energy terms and to construct energy potentials that take into account classical and non-classical terms. In the next sections, we will examine the dependency of these terms with respect to bond elongation and compare them with previous potentials.

The original IQA implementation could only deal with HF wavefunctions; however, recent developments provide support for DFT-derived ones.¹⁰ IQA calculations were performed with PROMOLDEN.^{9,33}

Results and Discussion

1. The model

As a starting point, the ELF-IQA approach was applied to a series of simple yet representative molecules in the $\text{CH}_3\text{-X}$ ($\text{X} = \text{CH}_3$, NH_2 , OH , F) series. In order to analyse the variation of the different terms that contribute to E_{intra}^{Bond} and E_{inter}^{A-Bond} (A-Bond), and attempt to rationalize the observed behaviour; the C-X interatomic distance, R , was stretched around the equilibrium position, R_{eq} , approximately in the range $[0.90\text{--}1.10] \cdot R_{eq}$. Since

we are resorting to small elongations, the use of a DFT approach is justified. The following standard ELF notation is used: $V(\text{C,X})$ represents the C-X bond, $V(\text{X})$ the lone pairs on X, $V(\text{C,Y})$ or simply Y, designates the core of atom Y and so on. For the sake of clarity, we will only analyse the results obtained for $\text{CH}_3\text{-NH}_2$ in detail, as an illustrative example with bonds and lone pairs. Similar results have been obtained for all the molecules in the $\text{CH}_3\text{-X}$ test set, and are provided in the Supporting Information.

The $\text{CH}_3\text{-NH}_2$ ELF basins and their populations are shown in Figure 1a. The evolution of the ELF populations of $V(\text{N})$ and $V(\text{C,N})$ upon bond stretching follows the expected behaviour, with a flux of electrons towards the lone pairs as the distance increases (see Figure 1b). In order to develop working equations along the stretching, we need to understand the evolution of the main parameters during this process. Bond charges and bonding basin sizes were found to change linearly with R (as shown in Figures 1b and 1c). Similar results for the complete $\text{CH}_3\text{-X}$ set are provided in Figures S1 and S2.

In the following sections, the interacting potentials for the kinetic, electrostatic and exchange correlation terms involving the C-X bond and valence lone pairs will be developed from the ELF-IQA partition. In all cases, the difference between the calculated IQA energy and the DFT obtained one (ΔE) was used to evaluate the numerical integration accuracy. The error held roughly constant and lower than 1% of the total energy, as shown in Table S7.

Classical electrostatic interaction energy. Since the ELF basins are non-overlapping, the electrostatic interaction between them can be reasonably well approximated by the Coulombic interaction between point charges located at the barycenter of the basin charge density. We have carried out a test of the performance of such an approximation in the reference molecule, $\text{CH}_3\text{-NH}_2$. The centre of charges for the C-N bond only deviates 0.07 a.u. from the ELF attractor. This small difference is clearly negligible when compared to the bond length, 2.75 a.u. Hence, monopole electrostatics are expected

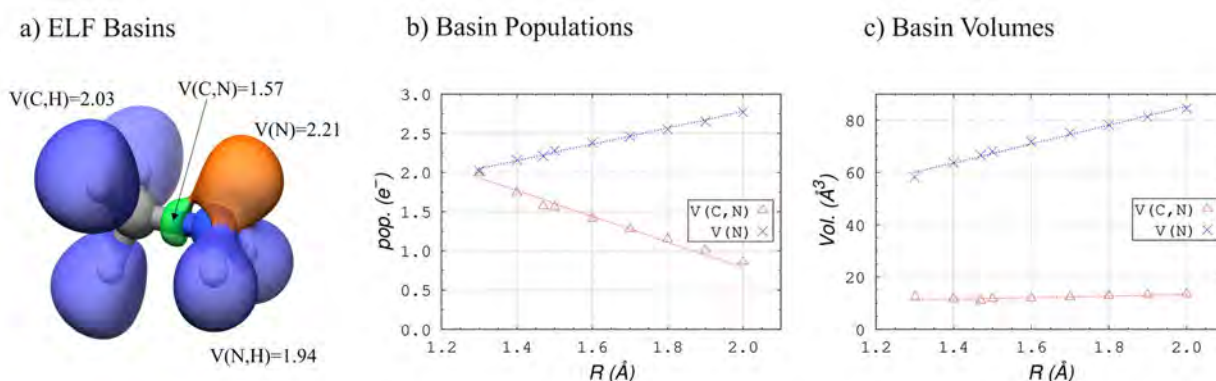


Figure 1. a) $\text{CH}_3\text{-X}$ ELF basins (isovalue = 0.8) and populations (at the equilibrium geometry). $V(\text{C,H})$ basins are shown in blue, $V(\text{C,C})$ and $V(\text{C,N})$ in green, and $V(\text{N})$ in orange. b) $V(\text{C,N})$ and $V(\text{N})$ populations as a function of the C-N bond distance. Linear regression coefficients: $R^2 = 0.982$ for $V(\text{C,N})$ and 0.989 for $V(\text{N})$. c) $V(\text{C,N})$ and $V(\text{N})$ volumes as a function of the C-N bond distance. Linear regression coefficients: $R^2 = 0.631$ for $V(\text{C,N})$ and 0.988 for $V(\text{N})$.

to reproduce reasonably well the electrostatic interactions (see eq. 8).

$$E_{elec}^{AB} \propto \frac{q^A \cdot q^B}{R} \quad (8)$$

Where q^i refers to the electronic charge of basin i (be it core, bond or lone pair). Hence, the overall electrostatic interaction should follow the classical interacting potential $E_{elec} = B/R + C$, where B and C are constants. For testing this hypothesis, we used the ELF-IQA approach to calculate the classical component of E_{intra}^{Bond} and E_{inter}^{A-Bond} (see eq. 6 and 7) for the interactions between the C–X basin (V(C,X)) and both the C and X (C(C) and C(X)) ones. In addition, we computed the interactions between the X lone pairs (V(X)) and X (C(X)). As previously explained, these interaction between basins represents the interaction between the C–X bond and C and X atoms, as well as the interactions between the lone pair in X and the X core, respectively. We represented these contributions against $1/R$. Results for the interaction between the pairs of valence basins in methanamine are shown in Figure 2 (see Figure S4 for the results involving all other molecules). As it can be seen, the linear regression fittings present very good correlation

coefficients, with R^2 being higher than 0.97 in all cases, for both intra and inter-basin contributions. The good agreement confirms that the ELF maximum can be used to approximate a point charge behavior. Remarkably, the linear correlation between the electrostatic interactions and $1/R$ also holds for intra-basin interactions, where $q_A = q_B$ (*i.e.* self-interaction can be neglected in the energy-model). Note that repulsive interactions are obtained for intra-basin terms, since they correspond to electron-electron repulsions, whereas inter-basin terms are positive, due to the extra attractive contribution from the nucleus. This agreement is also maintained when the basin charges (*i.e.* the electron density integrated over the basin volume) are taken into account (see Figure S5). Overall, this indicates that zero-th order electrostatics is a reasonably accurate approach for ELF basins.

Kinetic energy. In order to analyse the evolution of the valence electron pairs kinetic energy, we can assume that the electron density in the low density region (*i.e.* bonds far from the core) can be described by a zero-th order expansion of the kinetic energy density of a homogeneous electron gas, $t(r) = c_F \rho^{5/3}$. Assuming now a linear behavior of charges upon stretching (as shown in Figures 1b and S2), we obtain a simple expression relating the kinetic energy of valence basins with distance (see eq. 9). The development of eq. 9 is provided in the Supporting Information.

$$T = \frac{4\pi q^{5/3}}{3} \frac{1}{R_B^2} \quad (9)$$

The expression provided in eq. 9 depends on the bond basin radius (R_B). In order to verify whether we can relate it with the C–X bond distance, we calculate R_B as a function of R . For that, we subtracted the core radii to the interatomic distance, as proposed by Savin et al.³⁴ The results show that R_B (and R_{LP}) are proportional to R (see Figure S3). Therefore, we can develop a working expression of the form: $T \propto R^{-2}$. The validity of this relation for the kinetic energy of the valence basins of $\text{CH}_3\text{-NH}_2$, V(C,N) and V(N), is proven in Figure 3. The fittings for other molecules are provided in the Supporting Information (Figure S6). Overall, the agreement within the whole series of

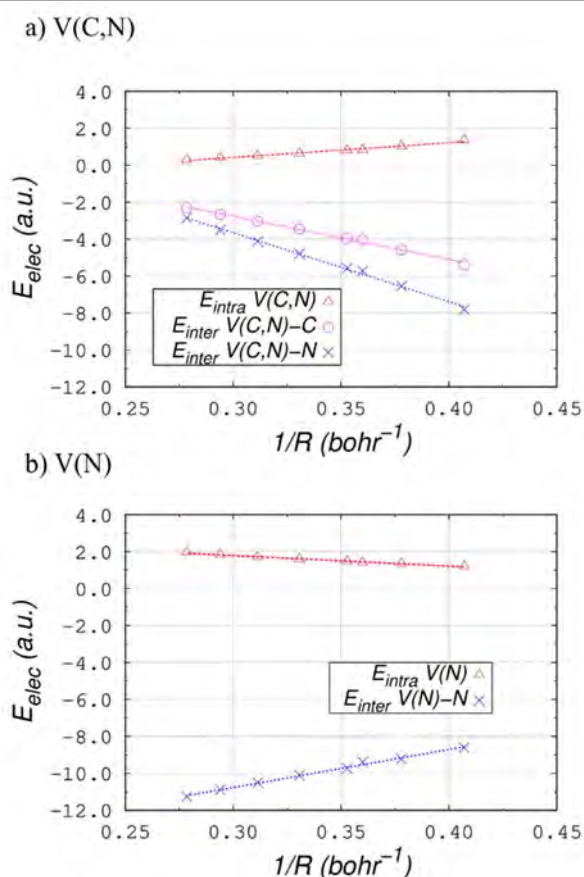


Figure 2. ELF-IQA computed classical electrostatic interaction energy (intra and inter) against $1/R$ for a) V(C,N) and, b) V(N) in $\text{CH}_3\text{-NH}_2$. Linear regression coefficients: $R^2 = 0.978$ for E_{intra} V(C,N); 0.995 for E_{inter} V(C,N)-C; 0.996 for E_{inter} V(C,N)-N; 0.989 for E_{intra} V(N); and 0.994 for E_{inter} V(N)-N.

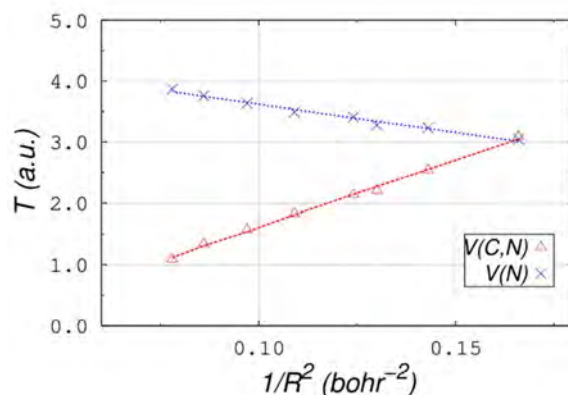


Figure 3. ELF-IQA computed bond kinetic energy against $1/R^2$ for V(C,N) and V(N) basins in $\text{CH}_3\text{-NH}_2$. Linear regression coefficients: $R^2 = 0.997$ for V(C,N) and 0.984 for V(N).

molecules leads to regression coefficients, R^2 , ranging from 0.984 to 1.00 for both, bonds and lone pairs.

Exchange-Correlation energy. Finally, let us examine the quantum mechanical contributions. Exchange-Correlation interactions have been shown to follow a linear behavior with the delocalization index, δ^{AB} , according to eq. 10.³⁵

$$E_{x-c}^{AB} \approx -\frac{\delta^{AB}}{2R} \quad (10)$$

Delocalization indices between QTAIM atoms have been shown to decay exponentially with the inter-centre distance for insulators,³⁶ as they depend on orbital overlap. In the case of ELF partitions, we have found that δ^{AB} has a clear hyperbolic behavior (that is, $\delta^{AB} = k/R$) for interactions between the bond and core basins in the range of distances considered (as shown in Figure S7). Accordingly, the effective expression for the exchange-correlation energy should follow eq. 11. Figure 4 shows how well this relation holds for the amine, the fittings for all the other molecules being provided in Figure S8 of the Supporting Information. Very good linear regression coefficients, $R^2 > 0.99$ in most cases, were obtained, confirming the validity of the previous reasoning and the derived dependencies.

$$E_{x-c}^{AB} \approx -\frac{\delta^{AB}}{2R} \approx -\frac{k'}{R^2} \quad (11)$$

Interplay of terms. The previous results have important consequences that are highly relevant in both, force field development, and quantum chemical topology. Firstly, they confirm the ability of an electron pair force field (that includes chemical bonds and lone pairs) derived from topology to describe a system while using well-behaved quantum contributions. The latter are easy to parametrize and to be added to available "classical" potentials.

Secondly, the quantum mechanical exchange-correlation energy term is found to be noticeably smaller than the classical terms for all systems in the dataset (see Figure 5). As an

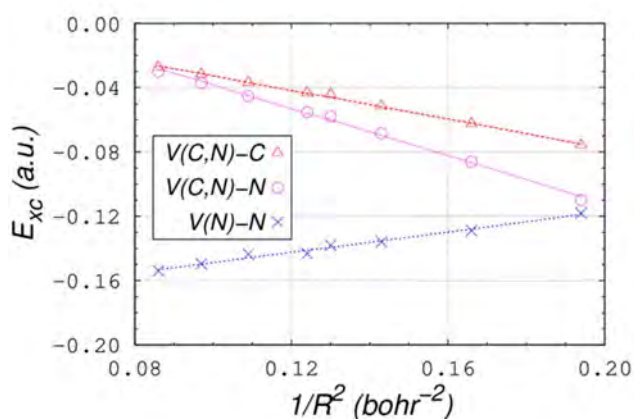


Figure 4. ELF-IQA computed exchange-correlation energies against $1/R^2$ for $V(C,N)$ and $V(N)$ basins with C and N in CH_3-NH_2 . Linear regression coefficients: $R^2 = 0.998$ for $V(C,N)-C$; 0.997 for $V(C,N)-N$; and 0.988 for $V(N)-N$.

example, we provide the different contributions for methanamine in the equilibrium position. For the $V(C,N)$ basins: $T_{V(C,N)} = 2.207$ a.u., $E_{elec-intra V(C,N)} = 0.825$ a.u., $E_{elec-inter V(C,N)-C} = -4.028$ a.u., $E_{elec-inter V(C,N)-N} = -5.723$ a.u., $E_{x-c V(C,N)-C} = -0.044$ a.u., $E_{x-c V(C,N)-N} = -0.058$ a.u. For $V(N)$: $T_{V(N)} = 3.281$ a.u., $E_{elec-intra V(N)} = 1.401$ a.u., $E_{elec-inter V(N)-N} = -9.379$ a.u., $E_{x-c V(N)-N} = -0.138$ a.u.

It is interesting to note that exchange-correlation terms are two orders of magnitude smaller than classical terms for bond basins. Nonetheless, they are only one order of magnitude smaller in the case of lone pairs. This reveals the greater relevance of quantum effects in the description of lone pairs. When building an electron pair force field they can be neglected, in a first approximation, in the case of covalent bonds, but should be taken into account when developing accurate interacting potentials for lone pairs. Furthermore, this prevalence of quantum effects further justifies the explicit treatment of lone pairs in Force Fields, which are often simply included in atomic contributions.

As a final check, Table S8 shows that intra-basin core energies (kinetic, electrostatic and exchange-correlation) remain constant upon elongation. Hence, they can be ignored in the construction of the interacting potential.

2. Comparison with previous models

Up to now, we have coupled IQA with the ELF topology to obtain the different interaction energy terms involving Lewis-like entities (atoms, bonds and lone pairs). According to the previous discussion, the total bond energy would be the sum of kinetic, electrostatic and exchange-correlation contributions, even though in most cases the latter will play a minor role. Interestingly, this framework is formally equivalent to the one provided by the Bond Charge Model (BCM).⁶ This approach is a simple albeit chemically intuitive model for calculating the energy of molecules (or solids). In particular, a homonuclear diatomic molecule, say A_2 , is represented as two cores with $+q/2$ positive charge, each of them interacting through a bond

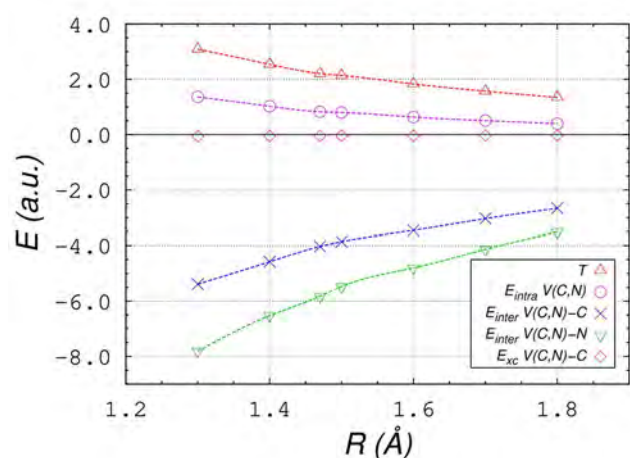


Figure 5. ELF-IQA energy terms for C-N bond in CH_3-NH_2 .

holding a charge of $-q$ electrons (see Figure 6 top). The bond charge is allowed to move along the bond length (R_B), which is a fraction of the equilibrium interatomic distance (R), $R_B = \nu R$. The total energy, which depends on the charge and the bond length, is the sum of three different contributions (see eq. 12, in atomic units).

$$E(q, R) = E_0 + E_1 + E_2 = 2E_A \left(\frac{q}{2}\right) - C \frac{q^2}{R} + D' \frac{q}{R_B^2} \quad (12)$$

Where E_0 is the core energy (equal to $2E_A$ for a homonuclear molecule, with E_A being the core energy of each atom); E_1 accounts for the classical electrostatic interactions (core-bond and core-core); and E_2 represents the kinetic energy of the bond electrons moving along the bond length, R_B . The model can also be expanded to heteronuclear systems, say AB, (see Figure 6 bottom) leading to eq. 13. This expression is formally equivalent to the one obtained for homonuclear molecules.

$$E(q, R) = E_A + E_B - C' \frac{q^2}{R} + D' \frac{q}{R_B^2} \quad (13)$$

Originally, the model parameters were obtained by fitting to experimental data.^{6,7} However, it is remarkable that the BCM model has the same energy terms that have been analysed by means of the ELF-IQA approach (with the exception of the exchange-correlation term, which is non-classical) and that they present the same dependencies with the interatomic distance, R . The parallelism with the terms derived from our ELF-IQA energy model is thus evident:

$$E_{core} = E_A$$

$$E_{intra}^{Bond} + E_{inter}^{A-Bond} = -C' \frac{q^2}{R}$$

$$T_{Bond} = D' \frac{q}{R_B^2}$$

$$E_{xc} \approx 0$$

The first three equations have been shown to work for our test set. As for the exchange-correlation contribution, a comparison of the scales of Figures 3 and 4 provides a first idea of the error. As previously stated, bond exchange-correlation terms are typically two orders of magnitude smaller than the other terms for bonds and one order of magnitude for lone pairs (see Figure 5).

It is important to note that this framework explains the previous success of the BCM model, as the energy terms with a classical origin are the only ones that contribute significantly to the scaling law of the total energy of molecules whenever electrons are well-localized. It also justifies the success of ELF-based parametrizations of the model. Moreover, it explains why this model was successful for extended solids in material science, while not so much applied for open structures (layers, chains), where lone pairs are usually present.³⁷ As we have seen, neglecting the exchange-correlation terms is straight forward for bonds (two orders of magnitude smaller than kinetic and electrostatic terms), but not so much for lone pairs (only one

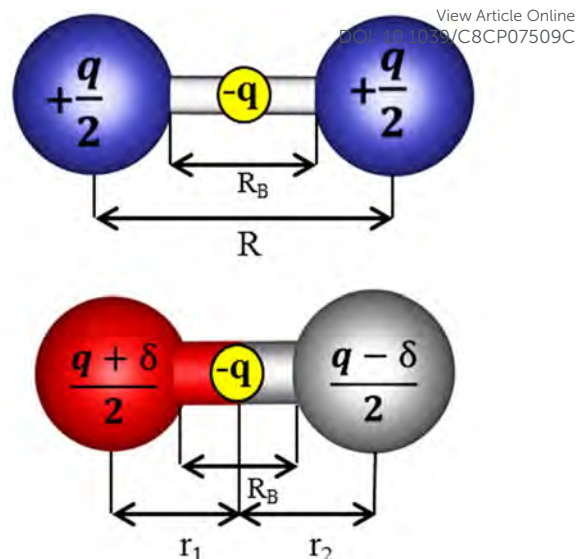


Figure 6. Bond Charge Model representation of a homonuclear molecule (top) and a heteronuclear one (bottom).

order of magnitude smaller). The next section will be devoted to the analysis of the limits of the IQA-ELF potential.

3. Limits of the model

Since the model is based on localized electron pairs, it is perfectly compatible with the description of covalent bonds. Accordingly, it may not suffice to capture other types of interactions. Several examples pertaining to distinct bonding regimes have been analysed as an attempt to estimate the range of applicability of the model. On the basis of the extent of electron pair localization, it should be possible to predict outlier cases where the IQA-ELF potential should fail. The following systems have been addressed: $\text{BH}_3\text{-NH}_3$, $\text{BeH}_2\text{-NH}_3$, $\text{CH}_3\text{-Li}$ and Li_2 , as paradigmatic examples of dative covalent bond, non-covalent beryllium bond, ionic and metallic systems (see Figure 7).^{38,39}

It was found that both, dative and very polar bonds, still follow the proposed dependencies for the IQA-ELF potentials. Specifically, the $\text{BH}_3\text{-NH}_3$ and $\text{BeH}_2\text{-NH}_3$ molecules lead to linear regression coefficients, R^2 , for the electrostatic and bond kinetic energies higher than 0.97 and 0.99, respectively (see Table S9 and S10). The exchange-correlation contribution was also found to be negligible in comparison with the classical energy terms, being two orders of magnitude lower, as in the $\text{CH}_3\text{-X}$ series. Similar results are obtained for $\text{CH}_3\text{-Li}$. Interestingly, the Li-V(C,Li) electrostatic term is smaller than the C-V(C,Li) counterpart by almost one order of magnitude. This result suggests that this contribution is not energetically meaningful, which can be related with the partially ionic character of the bond, which leads to the bonding basin being significantly closer to the carbon than to the lithium core (0.829 Å for C vs 1.169 Å for Li).

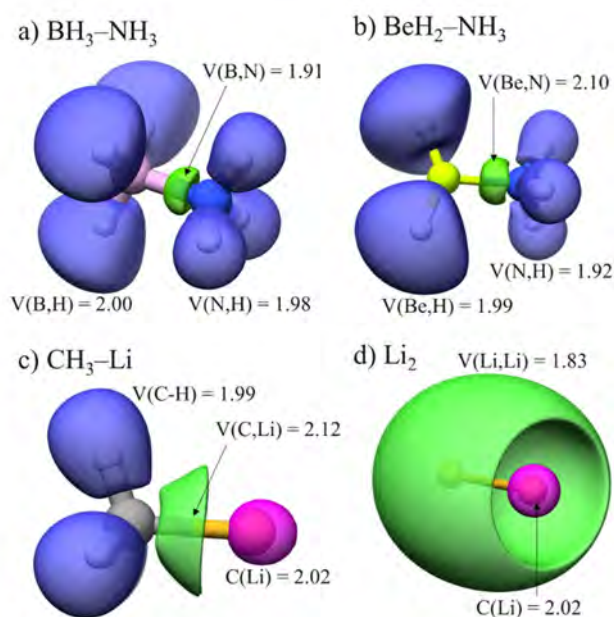


Figure 7. ELF basins (isovalue = 0.8) and populations for a) $\text{BH}_3\text{-NH}_3$, b) $\text{BeH}_2\text{-NH}_3$, c) $\text{CH}_3\text{-Li}$, and d) Li_2 . Hydrogenoid, disynaptic and core basins are shown in blue, green and purple respectively.

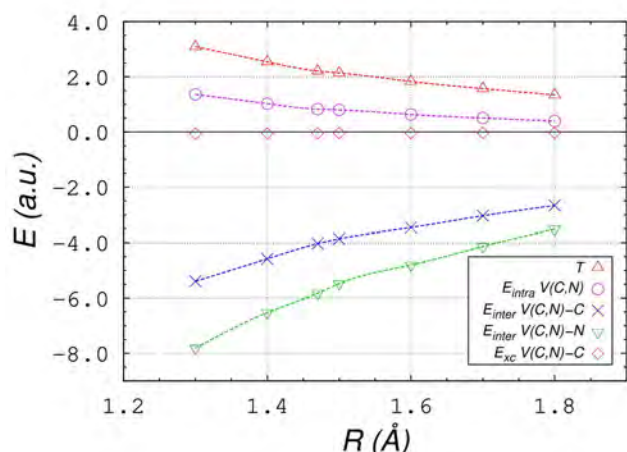


Figure 8. Kinetic, classical (intra and inter) and exchange-correlation energy terms for Li-Li bond in Li_2 . For simplicity, energy values are multiplied by 10^4 .

The proposed scaling laws of the energy terms do not hold for Li_2 . From the visual point of view, a rather remarkable feature is already identifiable in Figure 7d. The volume of the bond basin (where a non-nuclear maximum of the density appears) is much bigger than in previous cases (898.23 vs 17.91 \AA^3 in ethane). Moreover, the population in this basin also follows a different behavior upon stretching. It remains rather constant (between 1.81 and 1.83 e-) independently of the Li-Li interatomic distance. The resulting ELF-IQA kinetic, electrostatic and exchange-correlation energy terms are shown in Figure 8. It is noteworthy that all the terms are very small, of the order of 10^{-4} a.u., very different from the ones obtained for $\text{CH}_3\text{-X}$

molecules, which were of the order of 10^1 a.u. (kinetic and electrostatic terms) and 10^{-1} a.u. (exchange-correlation term). We can thus consider that the model works properly for all the molecules considered but Li_2 , for which it is not suitable.

Overall, the model is shown to be valid for covalent and partially covalent links, while failing for metallic bonds (as the case of Li_2). It should be noted that these two categories can be easily identified from properties derived from the electron density. Over the years, numerous attempts have been made to provide a distinct classification of the chemical bond from topological characteristics,^{40,41} with special emphasis on the metallic bond.⁴²⁻⁴⁴ Complex descriptors, as the metallicity index (ξ_m) proposed by Ayers *et al.*⁴⁴ have been used to further collect the character of bonds through topological measurements. This index is constructed to highlight regions of planar density, its expression being provided in eq. 14.

$$\xi_m(r_{bcp}) = \frac{36(3\pi^2)^{2/3} \rho(r_{bcp})^{5/3}}{5 \nabla^2 \rho(r_{bcp})} \quad (14)$$

High values of ξ_m are indicative of metallic character (typically $\xi_m > 25$ in solids); while weak metallic bonds are identified by ξ_m values between 1 and 5. When ξ_m is inferior to 1, the bond presents a non-metallic character. Bonds of all the previous molecules were characterized through their electron density, $\rho(r)$, its Laplacian, $\nabla^2 \rho(r)$, and the aforementioned derivative descriptor at the bond critical point. Given that local topological indices extrapolate the characteristics of a single point to a large chemical entity, a larger basis set has been used to guarantee converged densities and robustness. Results are presented in Table 1.

First of all, it is clear that the limiting cases we chose lead indeed to positive Laplacians and larger values of ξ_m . The metallicity index for Li_2 is much higher than for the rest of the considered bonds, which correlates with the breakdown of our model. Furthermore, the results do partially justify the lesser degree of accuracy of the model for the B-N, Be-N and C-Li bonds, as they exhibit greater ξ_m indices and positive Laplacian values at their BCP. This feature can be related to their only-partial covalent nature. All this is compatible with the basic idea encompassed here: localized pairs can be parametrized as point charges, so that when electrons indistinctly delocalize over a

Table 1. Topological characteristics of the selected bonds. Calculations were performed at the B3LYP/6-311G(d,p) level of theory.

Bond	ρ BCP (a.u)	$\nabla^2 \rho$ BCP (a.u)	$\xi_{m,BCP}^*$
C-C	0.238	-0.532	-5.518
C-N	0.261	-0.670	-5.116
C-O	0.256	-0.504	-6.560
C-F	0.233	-0.022	-130.6
B-N	0.099	0.420	1.631
Be-N	0.057	0.337	0.801
C-Li	0.044	0.213	0.821
Li-Li	0.013	0.002	12.78

* ξ_m is not defined when $\nabla^2 \rho < 0$, yet it can be calculated to exemplify the significant difference between bonds.

larger part of the system, the model fails. Contrarily, highly covalent bonds, with well localized pairs around their BCP can be assumed to be properly described by the model hereby presented.

Conclusions and outlook

We have presented a simple energy model based on the ELF topology for the stretching of bonds that explains the interaction between electron pairs while retaining their quantum mechanical character. Kinetic, electrostatic and exchange-correlation terms have been shown to follow distance dependencies that are akin to straightforward scaling laws. Electron pairs, as described by the Electron Localization Function, can be modelled from the homogeneous electron gas point of view. The proposed scaling behaviours have proven to work reasonably well for different polarities (from homonuclear to highly polar bonds, including non-conventional bonds). Whenever the electron pair fails to describe bonding, like in metallic bonds, or when highly delocalized bonding regimes set in, the model is expected to fail. We have also shown how simple local measures can be introduced to identify these situations.

Moreover, the energy model hereby introduced is able to explain the success of simple previous proposals, such as the Bond Charge Model, and its updated version, the ELF-BCM model.³⁷ In both of them, the quantum nature of the electrons was neglected. Our test set shows that, in general, the quantum mechanical contribution (represented by the exchange-correlation term) is much smaller than the other terms for bonds. On the contrary, quantum contributions become more important as electron delocalization increases, as in the case of lone pairs. Provided that lone pairs are of significant importance in many chemical systems that are usually modelled with Force Field approaches, we hereby justify the careful consideration of their position and electrostatic behaviour. In this regard, the definition of explicit lone pairs has already been considered in molecular mechanics modelling, either mathematically or conceptually.^{45,46}

Notably, the development of a molecular energy model such as the one here devised may have important implications in other closely related fields, such as conceptual DFT (c-DFT).⁴⁷ The existence of a simple parametrization for the energy of a system would allow to obtain molecular properties (electronegativity, chemical hardness, etc.) in a simple manner and starting from chemically meaningful concepts.^{48,49} As previously stated, another field on which this energy model may suppose a breakthrough is in Force Field development. In this respect, our model provides useful insight on non-classical terms and further justifies the design of Force Fields transcending the all-atom paradigm. Such Force Fields have already achieved remarkable success,⁵⁰ and we expect to foster these novel modelling methods. The work herein presented also opens the doors towards the further development of energy models based on Quantum Chemical Topology.

Conflicts of interest

There are no conflicts to declare.

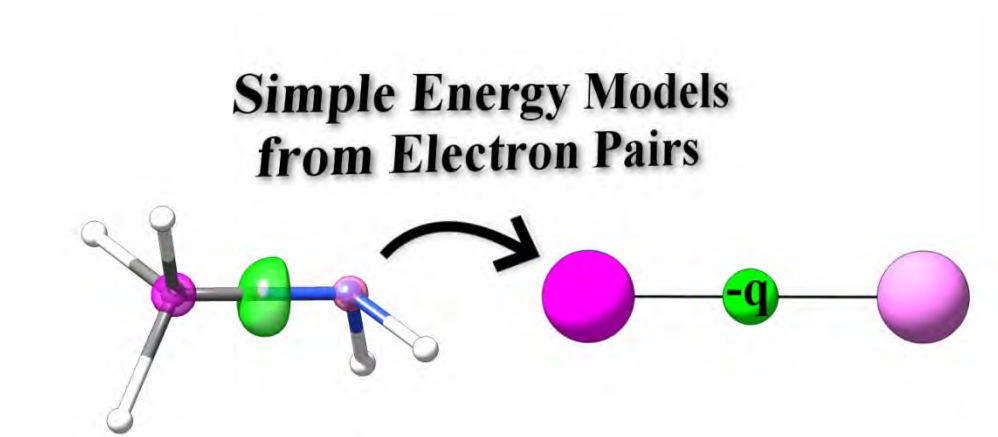
Acknowledgements

J.M. acknowledges financial support provided by the Spanish Ministerio de Ciencia, Innovación y Universidades (FPU14/06003 and EST17/00161), as well as from the Universidad de Zaragoza, Fundación Bancaria Ibercaja and Fundación CAI (CB 6/17). R.L. acknowledges ED388 for a PhD. grant. A.M.P. expresses its gratefulness for the financial support provided by the Spanish Ministerio de Ciencia, Innovación y Universidades (CTQ2015-65790-P). J.C.-G. acknowledges CALSIMLAB and the ANR within the Investissements d'Avenir program under reference ANR-11-IDEX-0004-02. In addition, the resources from the supercomputer "memento", as well as the technical expertise and assistance provided by BIFI-ZCAM (Universidad de Zaragoza) and the Laboratoire de Chimie Théorique (Sorbonne Université) are gratefully appreciated.

References

- N. L. Allinger, *Adv. Phys. Org. Chem.*, 1976, **13**, 1–82.
- J. H. McAliley and D. A. Bruce, *J. Chem. Theory Comput.*, 2011, **7**, 3756–3767.
- R. J. Gillespie and R. S. Nyholm, *Q. Rev. Chem. Soc.*, 1957, **11**, 339.
- S. Ekesan, S. Kale and J. Herzfeld, *J. Comput. Chem.*, 2014, **35**, 1159–1164.
- J. Herzfeld and S. Ekesan, *Phys. Chem. Chem. Phys.*, 2016, **18**, 30748–30753.
- R. G. Parr and R. F. Borkman, *J. Chem. Phys.*, 1968, **49**, 1055–1058.
- R. F. Borkman and R. G. Parr, *J. Chem. Phys.*, 1968, **48**, 1116–1126.
- M. A. Johnson, *Nat. Chem.* 2009, **1**, 8–9.
- A. M. Pendás, M. A. Blanco and E. Francisco, Two-electron integrations in the quantum theory of atoms in molecules, *J. Chem. Phys.*, 2004, **120**, 4581–4592.
- P. Maxwell, Á. M. Pendás and P. L. A. Popelier, *Phys. Chem. Chem. Phys.*, 2016, **18**, 20986–21000.
- O. A. Syzgantseva, V. Tognetti and L. Joubert, *J. Phys. Chem. A*, 2013, **117**, 8969–8980.
- J. M. Guevara-Vela, R. Chávez-Calvillo, M. García-Revilla, J. Hernández-Trujillo, O. Christiansen, E. Francisco, Á. Martín Pendás and T. Rocha-Rinza, *Chem. - A Eur. J.*, 2013, **19**, 14304–14315.
- J. Munarriz, E. Velez, M. A. Casado and V. Polo, *Phys. Chem. Chem. Phys.*, 2018, **20**, 1105–1113.
- P. L. A. Popelier, *Int. J. Quantum Chem.*, 2015, **115**, 1005–1011.
- N. Di Pasquale, S. J. Davie and P. L. A. Popelier, *J. Chem. Theory Comput.*, 2016, **12**, 1499–1513.
- F. Zielinski, P. I. Maxwell, T. L. Fletcher, S. J. Davie, N. Di Pasquale, S. Cardamone, M. J. L. Mills and P. L. A. Popelier, *Sci. Rep.*, 2017, **7**, 12817.
- A. Savin, R. Nesper, S. Wengert and T. F. Fässler, *Angew. Chemie Int. Ed. English*, 1997, **36**, 1808–1832.
- A. Martín Pendás, E. Francisco and M. A. Blanco, *Chem. Phys. Lett.*, 2008, **454**, 396–403.
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P.

- Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, 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, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision D.01*, Gaussian, Inc., Wallingford CT, 2009.
- 20 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
- 21 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 22 T. H. Dunning, *J. Chem. Phys.*, 1971, **55**, 716–723.
- 23 T. H. Dunning, K. A. Peterson and A. K. Wilson, *J. Chem. Phys.*, 2001, **114**, 9244–9253.
- 24 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- 25 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372–1377.
- 26 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158.
- 27 K. A. Peterson and T. H. Dunning, *J. Chem. Phys.*, 2002, **117**, 10548–10560.
- 28 A. Savin, O. Jepsen, J. Flad, O. K. Andersen, H. Preuss and H. G. von Schnering, *Angew. Chemie Int. Ed. English*, 1992, **31**, 187–188.
- 29 S. Noury, X. Krokidis, F. Fuster and B. Silvi, *Comput. Chem.*, 1999, **23**, 597–604.
- 30 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
- 31 T. D. Goddard, C. C. Huang and T. E. Ferrin, *J. Struct. Biol.*, 2007, **157**, 281–287.
- 32 M. A. Blanco, A. Martín Pendás, and E. Francisco, *J. Chem. Theory Comput.*, 2005, **1**, 1096–1109.
- 33 A. M. Pendás, E. Francisco and M. A. Blanco, *J. Comput. Chem.*, 2005, **26**, 344–351.
- 34 M. Kohout and A. Savin, *Int. J. Quantum Chem.*, 1996, **60**, 875–882.
- 35 E. Francisco, D. Menéndez Crespo, A. Costales and Á. Martín Pendás, *J. Comput. Chem.*, 2017, **38**, 816–829.
- 36 A. Gallo-Bueno, M. Kohout and A. Martín Pendás, *J. Chem. Theory Comput.*, 2016, **12**, 3053–3062.
- 37 J. Contreras-García, M. Marqués, J. Menéndez and J. Recio, *Int. J. Mol. Sci.*, 2015, **16**, 8151–8167.
- 38 K. M. Dreux, L. E. McNamara, J. T. Kelly, A. M. Wright, N. I. Hammer and G. S. Tschumper, *J. Phys. Chem. A*, 2017, **121**, 5884–5893.
- 39 M. Yáñez, P. Sanz, O. Mó, I. Alkorta and J. Elguero, *J. Chem. Theory Comput.*, 2009, **5**, 2763–2771.
- 40 D. Cremer and E. Kraka, *Angew. Chemie Int. Ed. English*, 1984, **23**, 627–628.
- 41 R. F. W. Bader and H. Essén, *J. Chem. Phys.*, 1984, **80**, 1943–1960.
- 42 P. Mori-Sánchez, A. M. Pendás and V. Luaña, *J. Am. Chem. Soc.*, 2002, **124**, 14721–14723.
- 43 G. Gervasio, R. Bianchi and D. Maraballo, *Chem. Phys. Lett.*, 2004, **387**, 481–484.
- 44 P. W. Ayers and S. Jenkins, *Comput. Theor. Chem.*, 2015, **1053**, 112–122.
- 45 J.-H. Lii and N. L. Allinger, *J. Phys. Chem. A*, 2008, **112**, 11903–11913.
- 46 T. Oroguchi and M. Nakasako, *Sci. Rep.*, 2017, **7**, 15859.
- 47 P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793–1874.
- 48 P. Geerlings and F. De Proft, *Phys. Chem. Chem. Phys.*, 2008, **10**, 3028–3042. View Article Online
DOI: 10.1039/C8CP07509C
- 49 R. Balawender, M. Lesiuk, F. De Proft and P. Geerlings, *J. Chem. Theory Comput.*, 2018, **14**, 1154–1168.
- 50 C. Bai and J. Herzfeld, *ACS Cent. Sci.*, 2016, **2**, 225–231.



465x200mm (72 x 72 DPI)

PCCP

Accepted Manuscript



This article can be cited before page numbers have been issued, to do this please use: J. Munarriz, R. Laplaza, Á. Martín Pendás and J. Contreras García, *Phys. Chem. Chem. Phys.*, 2019, DOI: 10.1039/C8CP07509C.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [author guidelines](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the ethical guidelines, outlined in our [author and reviewer resource centre](#), still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.