Communication: An approximation to Bader's topological atom

Pedro Salvador and Eloy Ramos-Cordoba

Articles you may be interested in

Communication: An approximation to Bader's topological atom

Pedro Salvador and Eloy Ramos-Cordoba[a)]

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

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

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

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

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

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

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

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

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

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

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

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

a)pedro.salvador@udg.edu

139, 071103-1

following step function

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

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

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

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

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

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

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

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

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

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

$$\mu_{AB} = \frac{r_A - r_B}{r_A + r_B}. \qquad (7)$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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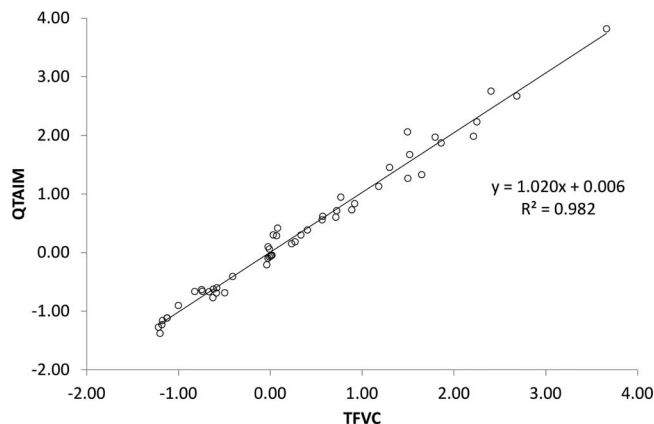

FIG. 1. QTAIM vs TFVC atomic charges for the molecular set.

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

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

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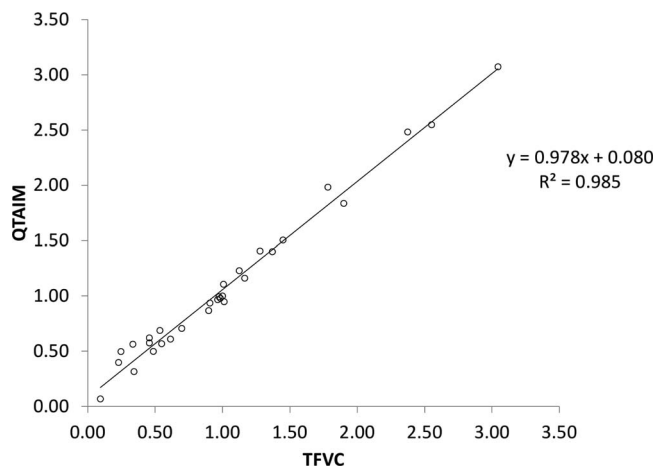

FIG. 2. QTAIM vs TFVC delocalization indices for the molecular set.

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

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

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

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

[1] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory* (Oxford University Press, Oxford, 1990).

[2] X. Fradera, M. A. Austen, and R. F. Bader, J. Phys. Chem. A **103**, 304 (1999).

[3] P. Salvador, M. Duran, and I. Mayer, J. Chem. Phys. **115**, 1153 (2001).

[4] M. A. Blanco, A. M. Pendás, and E. Francisco, J. Chem. Theory Comput. **1**, 1096 (2005).

[5] C. Gatti, Z. Kristallogr. **220**, 399 (2005).

[6] J. Friis, G. K. H. Madsen, F. K. Larsen, B. Jiang, K. Marthinsen, and R. Holmestad, J. Chem. Phys. **119**, 11359 (2003).

[7] A. M. Pendás, M. A. Blanco, A. Costales, P. Mori-Sánchez, and V. Luana, Phys. Rev. Lett. **83**, 1930 (1999).

[8] Q. K. Timerghazin, I. Rizvi, and G. H. Peslherbe, J. Phys. Chem. A **115**, 13201 (2011).

[9] D. Chesnut, Heteroat. Chem. **13**, 53 (2002).

[10] D. R. Alcoba, L. Lain, A. Torre, and R. C. Bochicchio, Chem. Phys. Lett. **407**, 379 (2005).

[11] I. Mayer and P. Salvador, Chem. Phys. Lett. **383**, 368 (2004).

[12] P. Salvador and I. Mayer, J. Phys. Chem. **120**, 5046 (2004).

[13] A. Becke, J. Chem. Phys. **88**, 2547 (1988).

[14] J. C. Slater, J. Chem. Phys. **41**, 3199 (1964).

[15] E. Matito, M. Solà, P. Salvador, and M. Duran, Faraday Discuss. **135**, 325 (2006).

[16] E. Francisco, A. M. Pendás, and M. A. Blanco, J. Chem. Theory Comput. **2**, 90 (2006).

[17] W. Heyndrickx, P. Salvador, P. Bultinck, M. Solà, and E. Matito, J. Comput. Chem. **32**, 386 (2011).

[18] See supplementary material at http://dx.doi.org/10.1063/1.4818751 for additional data such as integration accuracy of TFVC and use of the method with NNA's.

[19] P. Salvador and E. Ramos-Cordoba, "APOST-3D," University of Girona, Spain, 2013.

[20] J. Poater, X. Fradera, M. Duran, and M. Solà, Chem.-Eur. J. **9**, 400 (2003).

[21] P. M. Polestshuk, J. Comput. Chem. **34**, 206 (2013).

[22] M. Garcia-Borràs, S. Osuna, M. Swart, J. M. Luis, and M. Solà, "Maximum aromaticity as a guiding principle for the most suitable hosting cages in endohedral metallofullerenes," Angew. Chem. (published online).

[23] AIMAll (Version 13.05.06), Todd A. Keith, TK Gristmill Software, Overland Park KS, USA, 2013, see aim.tkgristmill.com.