New insights in quantum chemical topology studies using numerical grid-based analyses

David Kozlowski, Julien Pilmé

New Insights in Quantum Chemical Topology Studies Using Numerical Grid-based Analyses

David Kozlowski [1] and Julien Pilmé [2*]

1 Laboratoire de Chimie, UMR CNRS 5182, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. Present adress: CEA, DAM, DIF, F-91297 Arpajon, France

2 Laboratoire de Chimie Théorique, UMR 7616 CNRS, Université Pierre et Marie Curie, Sorbonne Universités, Case Courier 137, 4 place Jussieu, 75252 Paris Cedex 05, France.

*** Corresponding author:** pilme@lct.jussieu.fr

Fax: (33) 1 44 27 41 17

Abstract

We propose here new insights in Quantum Chemical Topology of one-electron density functions using a recent grid-based algorithm [*J. Phys. Condens. Matter.* **2009**, 21, 084204] initially designed for the decomposition of the electron density. Beyond the charge analysis, we show that this algorithm is suitable for different scalar functions showing a more complex topology, i.e., the Laplacian of the electron density, the electron localization function (ELF) and the Molecular Electrostatic potential (MEP). This algorithm makes use of a robust methodology enabling to numerically assign the data points of three-dimensional grids to basin volumes and it has the advantage of requiring only the values of the scalar function without details on the wave function used to build the grid. Our implementation is briefly outlined (program named TopChem), its capabilities are examined, and technical aspects in terms of cpu requirement and accuracy of the results are discussed. Illustrative examples for individual molecules and crystalline solids obtained with gaussian and plane-wave-based density functional theory calculations are presented. Special attention was given to the Molecular Electrostatic Potential (MEP) because its topological analysis is complex and scarce.

Keywords: topological analysis ; QCT; ELF ; MEP; QTAIM ; bonding analysis ; VASP.

1- Introduction

In recent years, several programs [1,2,3,4,5] devoted to the topological analysis of 3D scalar fields[6,7,8] (so-called Quantum Chemical Topology[9]) have been successfully designed for the analysis of chemical bonding or to provide a further understanding of the chemical reactivity. [10,11,12,13,14,15,16] Beyond the well-known approach of the "atoms in molecules" theory (QTAIM)[6,7] which relies on the properties of the electron density $\rho(\mathbf{r})$ when atoms interact, most of these programs are able to handle more complex topologies such as the topology of the Electron Localization Function (ELF).[17,18,19,20,21] In a topological analysis, the most important algorithm is the one that performs the assignment of the points on a three-dimensional grid to volumes (so-called basins) localized around the attractors (maxima) of the scalar function considered. These volumes are separated by surfaces (so-called separatrices). Some algorithms attempt to explore the accurate shape of these surfaces by an analytical procedure [22,23,24,25] but, this implies a large computational effort and may suffer from some convergence problems.[26] Thus, some programs such as TopMod[1], use partially a grid-based method in order to entails numerical difficulties and to reduce the computational time and effort. However, these codes require typically a wave function expressed in terms of cartesian gaussian primitives and related molecular orbital coefficients. Therefore, the computational cost is strongly linked to the numbers of molecular orbital coefficients and thereby rather restrict the applicability of programs to small molecular systems. Recently, the full numerical grid-based algorithms have evoked a renewed interest due to their robust efficiency and their relative easy implementations.[27,28] These numerical methods, where the scalar function is represented by its values on a regular grid, are suitable for a reasonable level accuracy calculations and are interesting for their reasonable computational cost which is mainly linked to the number of the grid points. Among these algorithms, Henkelman et al.[29,30,31] have introduced a new numerical grid-based method for the assignment of the points in the framework of the QTAIM methodology. This algorithm is solely based on the values of the electron density and performs the topological analysis of any grid of density, whatever the wave function initially used for building the grid. In addition, the algorithm does not attempt to give an explicit representation of separatrices which makes

the code very robust. Our purpose here, is to evaluate the applicability of this algorithm to any more elaborated topology (such as ELF grids for example). Another objective would be also to make possible the topological analysis of grids built from experimental determined electron densities.

2- Theoretical background

In this report, we assume that the reader is familiar with the topological analysis of scalar fields since several presentations of the method and possible applications have already been published in the literature. [7,10,32,33,34,35]. Briefly, the approach of the quantum chemical topology was devoted to provide an intuitive bridge between the traditional scheme of the chemical bond derived from the Lewis theory and first principles quantum-mechanical methodologies. In a topological analysis, a partitioning of the molecular space into subsystems (basins) is achieved by applying the theory of dynamical systems to the properties of the scalar function. The basins are localized around the attractors of the function and are separated by the zero flux surfaces (so-called separatrices). In the QTAIM theory, the function considered is the one electron density and the basins are associated with each of the atoms in the molecule. However, more elaborate functions have been used such as the Electronic Localization Function (ELF) of Becke and Edgecombe[17] which is interpreted as a signature of the electronic-pair distribution.[36,37] This function can be partitioned into an intuitive chemical scheme based on atomic and valence descriptors. The core basin volumes (if $Z > 2$) are located at nuclear centers. The valence basins (bonding and non-bonding regions) are situated in the remaining space and they are characterized by the number of core basins with which they share a common boundary. This number is called the synaptic order. [38] Monosynaptic basins (labeled V(A)) usually correspond to lone pair regions, whereas disynaptic and polysynaptic basins (labeled V(A, B, C, ..)) characterize the covalent bonds. Overall, the spatial distribution of the valence basins closely match the non-bonding and bonding domains of the VSEPR model. [39,40] The synaptic number is usually used for the ELF function in the framework of the scalar fields but it could be extended to other one-electron density function. Generally, the functions based on the density can exhibit topologies of various complexity. For instance, ELF is only positive, defined to have values between 0 and 1

and shows non atomic attractors. The Laplacian of the electron density ($\nabla^2\rho$), which may have either negative or positive values, displays also non-atomic attractors.

The basin populations can be then calculated by integrating the one electron density over the basin volume once the function was partitioned. A way to compute the basin polarization $\mathbf{M}_1$ (local dipolar moments) components was proposed in the framework of the QTAIM theory[41,42] and recently adapted to a more complex partition such as the ELF topology [43] (the basin is denoted by Ω) as follows:

$$M_{1x}(\Omega) = -\int_{\Omega} (x - X_c)\rho(\mathbf{r})\, d\tau$$

$$M_{1y}(\Omega) = -\int_{\Omega} (y - Y_c)\rho(\mathbf{r})\, d\tau$$

$$M_{1z}(\Omega) = -\int_{\Omega} (z - Z_c)\rho(\mathbf{r})\, d\tau$$

where X_c, Y_c, Z_c are the cartesian coordinates of the attractor (basin center). In this framework, the total molecular dipole (μ) is defined as the sum of contributions as follows: $\mu = \sum_{\Omega} M_1(\Omega) - Pop(\Omega) X_{\Omega} + Z.\mathbf{X}_{\Omega}$

where *Pop*(Ω) is the basin population, Z the atomic number and X_Ω is the location of the nuclear centers. Because of the invariance of the dipolar polarization magnitude (noted $|\mathbf{M}_1|$) with respect to the orientation of the system of axis, it is then possible to compare the polarization of bonds and lone pairs in different chemical environment. [43] This is a powerful tool to rationalize the intra and intermolecular interactions such as the hydrogen bond.

Numerical Grid-based Algorithm

The points of the grid are numerically attributed to core and valence basin volumes by means of the grid-based algorithm proposed by Henkelman et al. [29] Here no explicit representation of basin separatrices is determined and no attempt is undertaken to locate the saddle points. A cubic spline interpolation of the 3-dimensional grid was used to obtain interpolated gradient vectors at any location of the grid.[44]. Only the ascent trajectories confined to the grid points are followed (steepest ascent method) and used to

characterize the basin volumes, each grid point being considered only once. It is worth noting that the algorithm is very efficient and robust which makes the method applicable to complex topology such as the ELF one. In addition, this algorithm scales linearly with the number of the grid points. However, Sanville et al [30] have identified a bias with this algorithm which artificially moves the basin separatrices. To avoid this spurious behavior, the authors proposed a corrected algorithm[31] erasing the bias in which ascent trajectories along interpolated gradients were no more constrained to the grid. This corrected algorithm retains the linear scaling. A final refinement can be used to assign the wandering points located in the external molecular envelope. Indeed, in these regions, very low values are observed for the density functions and the gradient vectors are almost nul. This makes the assignment process very difficult. Thus, these wandering points are reassigned to the nearest neighbor points which have already an assignment code.

3- TopChem implementation and distribution

The grid-based corrected algorithm[31] and the calculation of the integrated quantities have been implemented in the TopChem program package, written in Fortran 90, containing two separated modules *chem_bas* and *chem_pop*. The overall organization of the package is built in the same spirit as the TopMod package previously developed in our laboratory. [1] Figure 1 provides the organization of TopChem. The *chem_bas* module requires only one input file in a correct cube format containing the values of the scalar function for each point of a rectangular parallelepipedic grid of dimension $N_x*N_y*N_z$. A routine implemented in *chem_bas* allows to convert the output files (CHGCAR, ELFCAR) generated by the Vienna Ab initio Simulation Package (VASP) into the required cube format. Three output files are created: a file giving the location of attractors on-grid and two files in binary format giving the number and the type of basin (i.e. core, bonding or non-bonding) for each grid point assigned. The *chem_pop* module provides integrated populations and local dipole moments. The module requires two input files: one input file in a correct cube format containing the values of the density function for each point and the output binary file of *chem_bas* giving the type of basin for each grid point assigned. In the TopChem implementation, the integrated quantities are simply

calculated by summing the electron density over the grid points assigned to the different basin volumes, without recourse to any analytical integration procedure. Of course, the accuracy of this method strongly depends of the number of grid points, however, we have shown that a step of 0.06 a.u. (see next section) between each grid point is quite sufficient to make the numerical procedure converge to the analytical results. Interestingly, the computational effort necessary to obtain integrated quantities remains very small in comparison to analytical integrations. In the end, the *chem_pop* module provides the basin volumes, the basin populations, the first moments and the related molecular dipole.

Distribution

The program written in Fortran 90 is available as freeware at : http://www.lct.jussieu.fr/pagesperso/pilme/topchempage.html. The tar file contains the Fortran sources, several Makefiles for 32bit and 64bit computer architectures under the Linux operating system and an user's manual.

An input and an output example are provided in a Supporting Information.

4- Performances and Applications

Computational methods

Geometry optimizations have been performed at the hybrid density functional B3LYP[45,46] level with the Gaussian 2003 software.[47] The MPWB1K hybrid meta-GGA functional[48] was employed for the geometry optimization of the methane hydrate structure because this functional is well-known for providing very good performances for hydrogen bonding and for weak interactions.[49] The empirically corrected for dispersion functional B97D[50], implemented in Gaussian09 [51], was also employed for the methane hydrate. The standard all-electron basis set cc-pVTZ was employed. The relativistic effective core potential (RECP) of the Stuttgart/Köln group was also used when appropriate.[52] In this case, only the valence electrons were explicitly taken into account, and the valence orbitals were described with the standard cc-pVTZ basis set. Results using the VASP program (Vienna Ab initio Simulation Package) [53] have been performed with the PW91 functional. [54]Core electrons are described by a PAW (projector augmented wave) procedure. [55] A plane wave basis set with a kinetic energy cut-off at

400 eV was used. The 3D-grids (ELF, density and its Laplacian) were written in the cube file either using the grid09 module[1] from a wfn gaussian output file or using the VASP output files (ELFCAR and CHGCAR). The grid of the MEP function was written in the cube file using the cubegen utility of Gaussian03 program. The isosurfaces were visualized with the Molekel software.[56]

Performances

To illustrate the applicability of our analysis, we present a series of results using a fine equidistant cube grid (step of 0.06 a.u.) for both charge density and ELF. These grids were built from different level of theory, i.e., DFT, HF, MP2, MP4SDQ and CISD. Table 1 presents the atomic charges and first moments of the oxygen atom in the water molecule for each level of theory. Our numerical results reproduce accurately the values obtained analytically using Gaussian expansion (Gaussian 03) taken as reference. For example, the oxygen atomic charge of -1.19e at the MP2/cc-pVTZ level compares favorably to the value of -1.18e obtained with Gaussian03. ELF populations in the water molecule have been also computed and compared satisfactorily to the related analytical references results provided by TopMod (see Table 1).

Scaling effort of the assignment algorithm

As noticed by Henkelman et al [29], the performance of the algorithm used for the assignment of the points depends mainly of the number of grid points. Indeed, Figure 2 displays the cpu time required to analyze the ELF cube files of several number of grid points for CH_3OH (6 atoms). The analysis time of grids ranging from 100 to 350 points was measured between 7s and 328s on a laptop. This Figure shows clearly that the computer time scales linearly with the number of grid points with a fixed computational effort per grid point. This result is a major key of our code to show its possible applicability to large molecular and solid state systems.

No computational effort is needed for the calculation of populations and local dipolar polarizations since they are computed by summing over the grid points assigned to basin volumes.

Influence of the grid density

Table 2 presents the QTAIM charges and the ELF populations for a single water molecule obtained from various sizes of a cubic grid. The influence of the spacing between adjacent grid points varying from a large step of 0.1 a.u (110^3 points) to a very fine step of 0.05 a.u (230^3 points) was analyzed. A medium grid size, i.e. a range between 0.1 a.u. and 0.08 a.u., provides quite reasonable results in comparison to accurate calculations obtained with TopMod (see Table 2). For example, the difference in the ELF population of V(O) basins using a very fine mesh of 0.05 a.u (230^3 points) and a medium mesh of 0.08 a.u. (140^3 points) was only found to be 0.05 electrons. Thus, because the computational effort is fixed per grid point, the study of large molecular systems becomes accessible using a medium grid size with a reasonable accuracy.

Using pseudopotentials within topological analysis

A special attention is needed if a pseudopotential representation of the core electrons is used. Indeed, no extrema of the local function is observed at the nuclear core and thus, only the outer-shell core (so-called subvalence) and valence regions are explicitly treated by TopChem. How our code deals with pseudopotentials was tested on CH_3OH, $Cu^I(H_2O)$ and XeN_2 where 2 core electrons of oxygen and carbon, 10 core electrons of copper, 46 core electrons (large core) or 28 core electrons (small core) of Xe were described by pseudopotentials whereas aug-cc-pVDZ basis set was used for the valence electrons. The ELF topological analysis of these three systems is given in Table 3. For CH_3OH, the populations of V(O) and V(C, O) have been respectively found to be 2.43e and of 1.29e which appears very close to the values obtained for the optimized system with an all-electron basis sets (see Table 3). However, some effects of the pseudopotential on the valence shells can be observed on the polarization of the oxygen lone pairs in CH_3OH. Indeed, the value of $|\mathbf{M}_1|$ (0.98 a.u) is enhanced in comparison to the value 0.85 a.u which has been calculated for the corresponding system using an all-electron basis sets. For $Cu^I(H_2O)$, the copper core basin C(Cu) displays a large population of 18.04e in perfect agreement with the likely oxidation degree of copper (+I) and the corresponding core population obtained with an optimized all-electron calculation (28.07e). For NXeN (C_{2v}), the small core representation provides an outer-shell basin

C(Xe) which shares separatrices with one nonbonding basin V(Xe) and two bonding basins V(Xe, N). The large core representation gives only one nonbonding valence basin V(Xe) and two bonding basins V(Xe, N). Similarly to previous molecules, the ELF topology of XeN_2 and the populations of valence basins appear quite similar to results obtained with TopMod (see Table 3). To conclude, all these investigations show the ability of TopChem to handle a pseudopotential core representation.

Applications.

We review here typical examples where the applicability of our analysis is illustrated for a large panel of molecular systems. We have mainly studied the topological analysis of the ELF function because of its complex topology (atomic and non atomic attractors) and, because its applicability to the chemical bonding analysis has been largely proved. [10] Another reason is the possible comparison of our results to the plentiful data that can be found in the literature. In addition, we present examples of analysis for more complex topologies in order to highlight the capabilities of the code to provide relevant populations. For this, we studied the topology of the negative regions of the Laplacian of the density (L(**r**) function) and the topology of the Molecular Electrostatic Potential (MEP).

Enabling direct insights into the Electron Localization Function

Table 4 reports the ELF population analysis of some bonds and lone pairs of interest for a various set of covalent molecules (H_2O, hydrocarbons and typical carbonyl systems), a dative system (NH_3BH_3), the neutral and zwitterionic forms of glycine, the H_2O_2 molecule and an example of d^0 metal compound ($CrMe_6$). The analysis of two open shell systems (CN, O_2) is presented as well. In addition, the analysis of the canonical water dimer is presented as a typical example of systems involving a hydrogen bond. Finally, the electronic structure of the methane hydrate structure is investigated. Results reported here are mainly focused on the populations and on dipolar polarization of covalent bonding basins and non-bonding pairs of interest. This study lead us to the following conclusions:

1. Overall, the calculated volumes and ELF populations as well as the calculated molecular dipoles are in excellent agreement with the previous data published in the literature.[32,43] A first example was reported with the bond populations of the V(C, C) basin of hydrocarbon systems which satisfactory increase along the sequence C_2H_6 (1.84e) , C_2H_4 (3.2e), C_6H_6 (2.77e) and C_2H_2 (5.06e) as a function of the formal C-C bond order. Another example can be found with BH_3NH_3 which is the archetype of the dative bond. When the dative bond N-B is formed, the number of basins is conserved but a non-bonding basin V(N) becomes a bonding basin V(B, N) with a population of 1.91e. In addition, the formation of the B-N bond does not yield any charge transfer between BH_3 and NH_3 moieties which differs strongly from an ionic interaction where no bonding basin is observed.

2. Another example is the canonical water dimer. As previously observed in a recent paper [43], the spatial distribution of the oxygen lone pairs V(O) can be used to characterize the hydrogen bond. Indeed, the dipolar polarization $|\mathbf{M}_1|$of V(O) involved in the H-bond (acceptor water) is enhanced (0.816 a.u.) by comparison to the other lone pair (0.765 a.u.). On the contrary, both lone pairs of the donor water appear similarly polarized with a value of 0.81 a.u. close to the values observed in the water monomer (0.78 a.u.). The origin of the polarization enhancement can be understood through a deep intramolecular reorganisation of the density within each water molecule due to the presence of the other molecule.

3. Some results for a family of related compounds involving a carbon–oxygen bond are presented. Table 4 gathers the volumes, the populations and $|\mathbf{M}_1|$ of the bonding basin V(C, O) together with those of the lone pairs V(O) for several typical C-O compounds, i.e., CH_3OH, CO_3^{2-}, H_2CO and CO. These molecules have been chosen according to their relatively different formal C-O Bond Order (BO). As expected, the populations and $|\mathbf{M}_1|$ values of the V(C, O) basin are clearly correlated to the bond order. For example, $|\mathbf{M}_1|$=0.075 a.u

for CH_3OH (BO=1) whereas $|\mathbf{M}_1|$=0.20 a.u for H_2CO(BO=2). We observe intermediates values for known systems exhibiting a π delocalization character such as CO_3^{2-}($|\mathbf{M}_1|$=0.563 a.u.). Our value of 0.183 a.u. for the dipolar polarization of the V(C, O) in the CO molecule indicates that the π contribution is weaker in comparison to other π systems such H_2CO ($|\mathbf{M}_1|$=0.20 a.u). This is consistent with recent studies which demonstrated that the π contributions and σ contributions are should be almost equivalent in the C-O bond of the carbon monoxide. [57]

The nature of O-O bond involved in the H_2O_2 molecule was briefly investigated by means of the ELF analysis. Indeed, this bond scheme is interesting since the O-O interaction is well-known as an archetype of the charge-shift bonding picture.[58] A disynaptic basin V(O, O) was found between the two oxygen atoms suggesting the presence of a O-O covalent bond. However, the V(O, O) basin displays a rather small population of 0.30 electrons (see Table 4), in comparison to the classical Lewis picture where the formal bond order is equal to 4 electrons. This small population is consistent with a charge-shift scheme where the O-O bond akin to a dissociated bond picture such as for the F-F bond in the F_2 molecule.

4. The ELF analysis of glycine is presented here because this molecule remains a subject of intensive study, for instance, in the field of the astrobiology. Indeed, enormous efforts for its detection in the interstellar medium were expended because this system is a symbol of the search for the origin of life in space[59,60,61] Briefly, several forms of this molecule can exist, depending mainly of the environment. In the gas phase, the neutral form NH_2CH_2COOH is stabilized while in a solvated or icy environments, the zwitterionic form $NH_3^+CH_2COO^-$ appears stabilized. We have studied the zwitterion with two different though related, models of solvation. On the one hand, $NH_3^+CH_2COO^-$ was only embedded in an electric continuum environment (PCM model[62] using the water dielectric constant ε = 80) and optimized at the B3LYP level of theory. On the other hand, the zwitterion was explicitly

surrounded by 50 water molecules in the same embedding PCM. This geometry of this system was optimized at the combined QM-QM level of theory (B3LYP for the high level of theory (glycine) and semi-empirical PM6 for the low level of theory (water molecules)). Then, a single point B3LYP has been made for the ELF analysis using the optimized QM-QM geometry. Table 4 gathers the ELF populations and moments while Figure 3 displays the ELF localization domain of glycine surrounded by the water molecules. Overall, the ELF analysis reveals an enhancement of the polarisations for the oxygen lone pairs in the zwitterion ($|\mathbf{M}_1|$=1.01 a.u) in comparison with its corresponding neutral form ($|\mathbf{M}_1|$=0.27 a.u). The two solvated models of glycine provide a quite similar response where the populations of V(C, O) bonding basins remain close to 2 electrons and the dipolar polarization of the oxygen lone pairs (V(O) basins) are almost unchanging ($|\mathbf{M}_1| \approx 1.0$ a.u) whatever the solvated model used (see Table 4). However, when the zwitterion is explicitly surrounded by 50 water molecules, it worth noting that a small charge transfer of 0.12e (see Table 4) is observed from the oxygen lone pairs to closest water molecules. The ELF analysis reveals here that although the embedding PCM cannot describe this charge transfer, it provides a quite reasonable description of the electronic structure of the polarized glycine in an embedded water environment.

5. An example of a CrX_6 compound (d0 molecule) has been analyzed. [63] The ideal VESPR geometry AX_6 is octahedral, as observed for CrF_6. However, a large number of such molecules have been found to have “non-VSEPR” geometries. In general, the deviations from the expected geometry can be explained by the crystal and ligand-field models as being due to the interaction of the ligands with the d electrons in the core. However, d0 complexes, formally at least, have no d electrons in the core. However, in a previous ELF study, [64] we have shown that the core basin has a population larger than the conventional value of 18e. The explanation can be found in the fact that the 3d basis functions centered on the metal noticeably contribute to

the electron density within the core region associated with the M shell. In addition, we have shown that under the influence of the ligands X, the metal core can lose its spherical symmetry by the formation of pairs localization basins (so-called subvalence), which in turn influence the geometry of the ligands if the interaction with the core is sufficiently strong to give non-VSEPR geometries. Here, the subvalence basins V(Cr) are the outer-shell core basins distributed around the central core basin C(Cr). It has been shown [10,64] that the location of the subvalence basins around the metal is related to the nature of the metal-ligand interaction (covalent or ionic). Such a case is observed here for $Cr(CH_3)_6$ (C_{3v} geometry) where the Cr-C interaction is strongly covalent. Our analysis shows that the ELF topology displays a core basin C(Cr) showing a large population of 21.15e, six bonding basins V(Cr, C) showing a rather large population of 1.25e each and six V(Cr) non-bonding basins (subvalence) distributed around the Cr core. All these results appear perfectly consistent with our previously published analysis using TopMod [64]

6. **ELF analysis of methane hydrate structure**.

Since the discovery in 1810 of the gas clathrate-hydrates, interest in hydrates has grown strongly. Indeed, crucial reasons make the contemporary study of clathrates a major challenge, which is at the same time energetic (they could be an important source of fossil energy) and environmental (agents of climate change). Clathrates are crystalline solids where water molecules form a cage-like structure around guest molecules, rather than in the hexagonal structure of normal ice. Different clathrate structures are known, the most common one being known as a structure where one guest molecule (typically methane) is incorporated in the cage formed of 20 molecules of water, as displayed in Figure 4. Overall, the structural parameters together with the stability aspects and the spectroscopic parameters of methane hydrate were already studied.[65,66]

The origin of the stabilization of the methane trapped in a water cage was studied by investigating a meaningful evaluation of the interactions between

the guest molecule and the cage by means of the ELF analysis implemented in TopChem. Our MPWB1K/6-31+G(d,p) optimized calculations show that the methane is stabilized by 2.01 kcal/mol[67] in a typical cage $(H_2O)_{20}$ (see Figure 4). This value is quite consistent with the stabilization of 5.0 kcal/mol obtained with the B97D functional for which dispersion was empirically corrected. All these results are also in good agreement with previous studies.[68] Table 4 reports the ELF analysis of the methane hydrate structure at the MPWB1K/6-31+G(d,p) level. No bonding basins were found between the methane and the cage. Thus, the nature of methane−cage interactions can be essentially described as electrostatic in which the water molecules polarize the methane subunit. Indeed, in comparison to the isolated methane (gas phase), the populations of C-H bonding basins V(C, H) increase slightly from 1.97e to 2.01e and their dipolar polarizations $|\mathbf{M}_1|$ increase from 1.0 a.u. to 1.05 a.u.. Thus, these results indicate that the methane is slightly polarized due to the environment of water molecules. This local induced polarization of the guest molecule due to the weak electrostatic interactions of water molecules appears directly related to the energetic stabilization of the methane in the cage.

Enabling direct insights into ELF and QTAIM populations obtained from plane waves expansions.

The need for a better understanding of larger and larger systems being crucial in today's research, we found it necessary to test applicability of our full numerical analysis for density and ELF grids issued from periodic descriptions of the solid matter (in this study provided by the VASP program). The ELF analysis in solids is scarce [69,70,71] because the periodicity gives rise to flat ELF profiles which entails difficulties for the assignment of points to the basins. However, the ELF topology of the core and valence regions are straightforward differentiated and our code allows to carry out the topological analysis of smoothly valence regions independently of the core regions where a pseudopotential can be used.

Four examples of systems are considered here. Figure 5 illustrates the ELF topology for benzene and the diamond lattice. Table 5 gathers the QTAIM analysis as

well as the ELF populations for four systems, i.e., C_2H_4 (sp2 carbon), C_6H_6 (aromatic), the crystalline diamond (sp3 carbon) and the ZnS crystal (ionic system). For C_2H_4, C_6H_6 and diamond, a bonding basin V(C, C) has been found with populations of 3.68e, 2.85e and 1.92e respectively. Thus, the bond multiplicity found for the C-C bond appears perfectly consistent with the nominal formal number of bonding electrons (4e for ethylene, 3e for benzene and 2e for C-C interaction in diamond lattice). The populations and the local dipoles $|\mathbf{M}_1|$ for C_2H_4 and C_6H_6 are also in agreement with related values obtained with TopMod (see Table 5). In addition, the ELF populations of the ZnS ionic crystal are presented in Table 5. Core basins C(Zn) and C(S) as well as valence basins V(Zn) and V(S) are found. As expected for an ionic interaction, the Zn-S interaction does not show any bonding basin V(Zn, S) between the cation Zn^{2+} and the anion S^{2-}.

Enabling direct insights into the Laplacian of the density and MEP functions.

The sign of $\nabla^2\rho$ determines whether the density is concentrated ($\nabla^2\rho < 0$) or depleted ($\nabla^2\rho > 0$) at a point **r** relatively to the surrounding points. Thus, the topology of the opposite function $L(\mathbf{r}) = -\nabla^2\rho$ is interesting in the positive regions because it indicates the concentrated density corresponding to core, bonds and lone pairs regions. Some attempts to perform its topological analysis have already been reported in previous studies.[72,73] The applicability of our grid-based analysis is shown here on the water molecule. Table 6 presents the population analysis and Figure 6 displays the localization domains of L(r). This topology appears similar to ELF except for the H atoms where an additional core basin is found (see Figure 6). Similarly to ELF, the valence basins closely match the domains of the VSEPR model. Indeed, two maxima for the O-H bonds and two maxima for the oxygen lone pairs were found. These maxima are distributed in a tetrahedral arrangement centered on the oxygen atom.[7,72] The synapticity, usually applied to ELF, can be extended to the L(r) function where each valence basin shares a common boundary with one or several cores.[72] For example, each bonding basin O-H displays a disynaptic character V(O, H) because the basin shares one boundary with the oxygen core C(O) and shares another boundary with the hydrogen core C(H). However, the basin populations differ strongly from ELF because the lone pair has a population of 1.68e, the core of oxygen shows a population of 0.83e and the two O-H bonds displays a population

of 0.57e. The unexpectedly small populations of O-H bonds and for the oxygen core, nearer to one than to two electrons, are in reasonable agreement with a previous published study.[72]

Another interesting function is the Molecular Electrostatic Potential (MEP). From a fundamental point of view, the MEP is an observable and the location of its critical points hold a key to understanding the reactivity of molecules. [74,75,76,77] The topology of MEP is much complex because the function can take negative or positive values. In the positive regions, the MEP topology shows solely local maxima at the atomic sites. However, in the negative regions non nuclear maxima occur in its topology distributed generally in agreement with the location of VSEPR domains. The topological analysis of the MEP was carried out for each positive and negative region.The localization domains of each regions were merged in Figure 6. Volumes and populations are gathered in Table 6. Core basins for oxygen and hydrogen atoms have been identified in the MEP positive part whereas two maxima for oxygen lone pairs have been found in the negative regions. Contrary to the populations of L(r), the oxygen core displays a large population of 7.58e while the populations of oxygen lone pairs are unexpectedly small to be 0.33 electrons each. In agreement with a published previous study, [74] the distance of attractors associated to lone pairs from the oxygen atom has been found very large (1.24 Å) in comparison to the distance of the same attractor for the ELF function (0.59 Å) This, probably explains the small populations of the lone pairs since the values of the electron density strongly decreases away from the oxygen atom.

These two examples of the complex analysis illustrate the capabilities of numerical analysis to handle different one-electron density functions. Our future goal will be to expand its capabilities to other functions of interest for the chemical bond or in reactivity such as the Dual Descriptor[78] which can be defined as the difference between nucleophilic and electrophilic Fukui functions.

5- Conclusions

The evaluation of an efficient and robust algorithm able to perform a topological analysis of elaborated one-electron density functions and its implementation were presented. Our code implementation in the TopChem program has been specifically

designed to make the analysis possible regardless of the method or the quantum chemistry program used to generate the grid points. In light of these investigations, the ability of the new algorithm to perform topological analysis of different scalar functions showing a more or less complex topology ($\rho(\mathbf{r})$, $-\nabla^2\rho$, ELF, MEP) was demonstrated.

Future works will focus on the development of new potentialities to allow the analysis of other functions of interest for the bonding analysis such as the Electron Pair Localization Function (EPLF) [79] or other functions of interest in chemical reactivity such as the Dual Descriptor.[78] In addition, expanding the applicability to ELF grids implemented in other quantum chemistry program as the Car-Parrinello Molecular Dynamics package (CPMD) [80] is currently under development.

6- Acknowledgments

The authors would like to thank Yves Ellinger (LCT) and Paul Fleurat-Lessard (ENS Lyon) for their critical reading of the manuscript and for stimulating discussions. Parts of the calculations presented in this contribution were financed by the "Centre Informatique National de l'Enseignement Supérieur" (CINES) supercomputing center and the "Pôle Scientifique de Modélisation Numérique" (PSMN) at the Ecole Normale Supérieure de Lyon.

References

[1] Noury, S.; Krokidis, X.; Fuster, F; Silvi, B. *Comp & Chem.* **1999**, 23, 597.

[2] Popelier, P.L.A. *Comp. Phys. Comm.* **1996**, 93, 212-240.

[3] Kohout, M. Chemical Bonding Analysis in Direct Space **2002**. http://www.cpfs.mpg.de/ELF.

[4] Biegler-König, F; Schönbohm, J.; Bayles, D. *J. Comput. Chem.* **2001**, 22, 545-559.

[5] Otero-de-la-Roza, M.A. ; Blanco, A. ; Pendás, M ; Luaña V. *Comp. Phys. Comm.* **2009**, 180, 1, 157-166.

[6] Bader, R. F. W. "Atoms in Molecules: A Quantum Theory". Oxford Univ. Press., Oxford, **1990**.

[7] Gillespie, R. J. ; Popelier, P. L. A. Chemical Bonding and Molecular Geometry, Oxford University Press, Oxford, **2000**

[8] Silvi, B. ; Savin A. *Nature* **1994**, 371, 683.

[9] Popelier, P. ; Ed. D.J.Wales, *Structure and Bonding*, Springer-Verlag, **2005**, 115, 1-56 .

[10] de Courcy, B. ; Pedersen, L. G. ; Parisel, O.; Gresh, N.; Silvi, B.; Pilmé, J. and Piquemal, J-P *J. Chem. Theory. Comput.* **2010**, 6, 4, 1048–1063.

[11] Pilmé, J. ; Bertoumieux, H. ; Robert V. and Fleurat-Lessard, P. ; *Chem. Eur. J.* **2007**, 13, 19, 5388-5393.

[12] de la Lande, A. ; Salahub, D. R. ; Maddaluno, J. ; Scemama, A ; Pilmé , J. ; Parisel, O. ; Gérard, H. ; Caffarel, M. ; Piquemal, J-P *J. of Comput. Chem.* **2010** 32, 6, 1178-1182.

[13] Berski, S. ; Andrés, J; Silvi, B.; Domingo, L. R.; *J. Phys. Chem. A* **2006;** 110, 13939.

[14] Poater, J. ; Duran, M.; Solà, M.; Silvi, B.; *Chem. Rev.* **2005**, 105, 3911.

[15] Pauzat, F.; Pilmé, J. ; Toulouse, J. ; Ellinger, Y.; Salama, F. *J. Chem. Phys*. **2010**, 133, 054301.

[16] Rivera-Fuentes, P. ; Aonso-Gomez, J. L. ; Petrovic, A.G.; Seiler, P.; Santoro, F. ; Harada, N.; Berova, N ; Rzepa, H. S. ; Diederich, F. ; *Chem. Eur. J.* **2010** 16,32 ,9796-9807.

[17] Becke, D. ;. Edgecombe, K. E. *J. Chem. Phys.* **1990**, 92, 5397.

[18] Savin, A.; Nesper, R.; Wengert, S.; Fässler, T. F.; *Angew. Chem. Int. Ed. Eng*. **1997**, 36, 1809.

[19] Silvi, B. *J. Phys. Chem. A* **2003**, 107, 3081.

[20] Matito, E.; Silvi, B.; Duran, M.; Solà, M. *J Chem Phys.* **2006**, 125, 2, 24301.

[21] Feixas, F.; Matito, E. ;Duran, M. ; Sola, M. ; Silvi B. ; *J. Chem. Theory Comput.* **2010**, 6, 9, 2736–2742.

[22] Popelier, P.L.A. *Theor. Chim. Acta*. **1994**, 87, 465.

[23] Popelier, P.L.A. *Chem. Phys. Letters.* **1994**, 228, 1-3, 160-164

[24] Popelier, P.L.A. *Comp. Phys. Comm.* **1998**, 108, 180-190

[25] Malcolm, N.O.J. ; Popelier, P.L.A. *J. Comput. Chem.* **2003**, 24, 437-442

[26] Uberuaga, B. P. ; Batista, E.R. ; Jónsson, H. *J. Chem. Phys.* **1999**, 111,10664

[27] Rodríguez, J. I. ;Köster, A. M.; Ayers, P. W.; Santos-Valle, A.; Vela, A.; Merino, G. *J. Comput. Chem.* **2009**, 30, 7, 1082-1092

[28] Rodríguez, J. I.; Bader, R. F. W.; Ayers, P. W.; Michel, C.; Götz, A. W.; Bo, C. *Chemical Physics Letters* **2009**, 472, 149–152

[29] Henkelman, G; Arnaldsson, A; Jónsson, H. *Comput. Mater. Sci.* **2006**, 36, 254-360.

[30] Sanville, E. ; Kenny, S. D. ; Smith, R.; Henkelman, G *J. Comp. Chem.* **2007**, 28, 899-908

[31] Tang, W.; Sanville, E; Henkelman, G. *J. Phys. Condens. Matter.* **2009**, 21, 084204.

[32] Silvi, B; Fourre, I.; Alikhani, M. E. ; *Monatshefte Für Chemie* **2005**, 136, 6, 855-879.

[33] Contreras-Garcia, J ; Recio, J. M. *J. Phys. Chem. C* **2011**, 115,1, 257-263.

[34] Piquemal, J.-P ; Pilmé, J. ; Parisel, O. ; Gérard, H. ; Fourré, I. ; Bergès, J. ; Gourlaouen, C. ; de la Lande, A. ; van Severen, M. C. ; Silvi, B. *Int. J. Quantum. Chem.* **2008**, 108,195.

[35] Silvi, B; Pilmé, J; Fuster, F; Alikhani, E. A.; In N. Russo, D. R. Salahub, and M. Witko (eds), "Metal-Ligand Interactions : Molecular-,Nano-Micro-Systems in Complex Environments". (Kluwer, Dordrecht), *NATO ASI SCIENCES SERIES: II : Mathematics, Physics and Chemistry*. **2003** 116, 241-284.

[36] Feixas, F.; Matito, E.; Duran, M., Sola, M.; Silvi, B. *J. Chem. Theory Comput.* **2010**, 6, 2736–2740.

[37] Silvi, B.; *J. Phys. Chem.* A **2003**, 107, 3081

[38] Silvi, B. ; *Journal of Molecular Structure*, **2002**, 614, 3–10

[39] Gillespie, R. J.; Nyholm R. S.; *Q. Rev. Chem. Soc*. **1957**, 11, 339.

[40] Gillespie, R. J.; *Molecular Geometry*, Van Nostrand Reinhold, London **1972**.

[41] Bedall, P.M.; Cade, P.E.; *J. Am. Chem. Soc.* **1971**, 93, 3095.

[42] Bader, R. F. W.; Larouche, A.; Gatti, C.; Caroll, M. T.; Dougall, P.J.; Wiberg, K. B. *J. Chem. Phys.* **1987**, 87,1142.

[43] Pilmé, J. , Piquemal, J.-P *J. Comput. Chem.* **2008**, 29, 9, 1440-1449

[44] Kenneth, P.; Esler, Jr. Spline libraries and documentation can be found at http://einspline.sourceforge.net/index.shtml

[45] Becke, A. D. *J. Chem. Phys* **1993**, 98, 5648.

[46] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, 37, 785.

[47] Gaussian 03, Revision B.02, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Pittsburgh PA, **2004**.

[48] Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, 120, 215.

[49] Dkhissi, A; Blossey, R; *Chemical Physics Letters* **2007**, 439, 1-3, 35-39

[50] Grimme, S; *J. Comp. Chem*., **2006**, 27, 1787-1799

[51] Frisch, M. J.Gaussian09, Rev. A.02; Gaussian Inc.: Wallingford CT, 2009

[52] Andrae, D.; Haeussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, 77, 123.

[53] Kresse, G. ; Hafner, J. ; *Phys. Rev. B* **1994**, 49, 14251.

[54] Perdew, J. P. ; Chevary, J. A. ; Vosko, S. H. ; Jackson, K. A. ; Pederson, M. R. ; Singh, D. J. ; Fiolhais, C. *Phys. Rev. B* **1992**, 46, 6671-87.

[55] Blöchl, P. E. *Phys. Rev. B* **1994**, 50, 17953.

[56] Flukiger, P.; Luthi, H. P.; Portman, S.; Weber, J. **2000-2002,** Molekel, Swiss Center for Scientific Computing, Manno, Switzerland

[57] Frenking, G.; Loschen, C; Krapp, A.; Fau; S.; Strauss, S. H.; *J. Comput. Chem.* **2006**, 28, 1,117

[58] Shaik, S; Danovich, D; Silvi, B.; Lauvergnat, D. L; Hiberty P. C. *Chem. Eur. J.* **2005**, 11, 6358 – 6371

[59] Kuan, Y.-J. ; Charnley, S. B. ; Huang, H.-C. *ApJ* **2003**, 593, 848.

[60] Jones, P. A.; Cunningham, M. R.; Godfrey, P. D.; Cragg, D. M. *MNRAS* **2007**, 374, 579

[61] Marloie, G. ; Lattelais, M. ; Pauzat, F. ; Ellinger, Y. *Interdiscip. Sci. Comput. Life Sci*. **2010**, 2, 1, 48-56.

[62] Tomasi, J. ; Persico, M. *Chem. Rev.* **1994**, 94, 2027

[63] Kaupp, M. *Angew. Chem., Int. Ed. Engl.* **2001**, 40, 3534.

[64] Gillespie, R. J. ; Noury, S. ; Pilmé, J. ; Silvi, B. *Inorg. Chem.* **2004**, 43, 3248-3256.

[65] Hermida-Ramón, J. M. ; Graña, A. M. ; Estévez, C. M. ; *Struct. Chem.* , **2007**, 18, 649–652

[66] Qi-Shi, D.; Da-Peng, L. ; Peng-Jun, Liu; Ri-Bo, Huang; *J. of Mol Graphics and Modelling*, **2008**, 27, 140–146

[67] The energy of stabilization of methane in cage is calculated as follows :
$E[CH_4(H_2O)_{20}] – E[(H_2O)_{20}] - E[CH_4]$. The three structures were optimized.

[68] Terleczky, P. ; Nyulászsian L. *Chem. Phys. Letters*, **2010**, 488, 168-172

[69] Savin, A; Jepsen, O. ; Flad, J. : Andersen, O. K. ; Preuss, H; von Schnering, H. G. *Angewandte Chemie International Edition in English.* **1999**, 31, 2, 187–188.

[70] Contreras-Garcia, J. ; Recio, J. M. *J. Phys. Chem. C* **2011**, 115, 257–263

[71] Contreras-Garcia, J.; Pendas, A. M.; Recio, J. M.; Silvi, B. *J. Chem. Theory Comput.* **2009**, 5, 1,164-173

[72] Malcolm, N. O. J. ; Popelier, P. L. A. *Faraday Discussions* **2003**, 124, 353-363.

[73] Popelier, P. L. A. *Coordination Chemistry Reviews* **2000** ,197, 169-189.

[74] Leboeuf, M ; Koster, A. M. ; Jug, K. ; Salahub, D. R. *J. Chem. Phys.* **1999**, 111, 11,4893-4905.

[75] Sayyed, F. B.; Suresh, C. H.; Gadre, S. R. *J . Phys. Chem A* **2010**, 114, 46, 12330-12333.

[76] Roy, D. K. ; Balanarayan, P. ; Gadre, S. R. *J. of. Chem. Sciences* **2009**, 121, 5, 815-821.

[77] Martinez, A.; Calaminici, P.; Koster, A. M. et al. *J. Chem. Phys.* **2001**, 114, 2, 819-825.

[78] Morell, C.; Grand, A.; Toro-Labbe, A. *J. Phys. Chem. A* **2005**, 109, 205-212

[79] Scemama, A.; Caffarel, M. ; Chaudret, R. ; Piquemal, J.-P *J. Chem. Theory Comput.* **2011**, 7, 3, 618.

[80] CPMD, http://www.cpmd.org/, Copyright IBM Corp **1990-2008**, Copyright MPI fur Festkorperforschung Stuttgart **1997-2001**.

Tables

H_2O	PW91	B3LYP	HF	MP2	MP4SDQ	CISD
QTAIM						
Oxygen charge	-1.11(-1.10)[a]	-1.12(-1.13)[a]	-1.24(-1.24)[a]	-1.19(-1.18)[a]	-1.16(-1.16)[a]	-1.18(-1.17)[a]
Oxygen dipole $\|\mathbf{M}_1\|$	0.28(0.26)[a]	0.28(0.24)[a]	0.38(0.39)[a]	0.35(0.35)[a]	0.34(0.32)[a]	0.34(0.33)[a]
ELF						
V(O) population	2.29(2.28)[b]	2.29(2.29)[b]	2.25(2.25)[b]	2.27(2.26)[b]	2.27(2.27)[b]	2.26(2.26)[b]
V(O, H) population	1.64(1.64)[b]	1.65(1.64)[b]	1.68(1.69)[b]	1.66(1.66)[b]	1.66(1.66)[b]	1.66(1.66)[b]

Table 1: QTAIM and ELF properties of the oxygen atom in the water molecule at different level of theory. The geometry of the water molecule was optimized at the B3LYP/cc-pVTZ level of theory. All topological analysis were performed using a grid step used of 0.06 a.u. a) Values in parenthesis are obtained with Gaussian03 where the QTAIM analysis is implemented. b) Values in parenthesis are obtained with TopMod.

step size (a.u.)	**0.10**	**0.09**	**0.08**	**0.07**	**0.06**	**0.05**
QTAIM						
Oxygen charge	-1.20	-1.17	-1.15	-1.14	-1.12	-1.12 (-1.12)[a]
Oxygen polarization $\|\mathbf{M}_1\|$	0.30	0.30	0.29	0.28	0.28	0.28 (0.28) [a]
ELF populations						
V(O)	2.21	2.23	2.24	2.28	2.29	2.29 (2.29) [a]
V(O, H)	1.64	1.66	1.64	1.65	1.65	1.64 (1.64) [a]

Table 2: Calculated charges and populations obtained with various spacing between adjacent grid points for a single water molecule optimized at the B3LYP/cc-pVTZ level of theory.
a) The values given in parenthesis are obtained with TopMod.

Molecular systems	Volume	Population	μ (D)[b]
CH_3OH			1.58 (1.61) [b]
V(C, O)	7.20(6.9)[c]	1.29 (1.24)[c]	
V(O)	46.1(49.7)[c]	2.43 (2.36)[c]	
Cu^IH_2O (C_{2v})			3.84 (3.82)[b]
C(Cu)	142.1 (149.1)[c]	18.04 (28.07)[c]	
V(O)	43.1 (42.6)[c]	2.21 (2.22)[c]	
XeN_2 (C_{2v})			
Large core (8 core elec.)			2.79 (2.81)[b]
V(Xe)	747	6.46 (6.43)[a]	
V(Xe, N)	272	2.79 (2.77)[a]	
Small core (26 core elec.)			2.90 (2.93)[b]
C(Xe)	22.2	17.84(17.80)[a]	
V(Xe)	743	6.35 (6.32)[a]	
V(Xe, N)	623	2.84 (2.85)[a]	

Table 3. Influence of the pseudopotential on the ELF populations.
Systems optimized at the B3LYP/aug-cc-pVDZ level of theory. The Stuttgart/Köln (SDD) core pseudopotential was used for C, O, Cu and Xe. The TopChem populations (in electrons) were obtained using a fine equidistant cube grid (step of 0.06 a.u.). The volumes (a.u.) were calculated with a threshold of 0.001 a.u. on the density.
a) In parenthesis, the values were obtained with TopMod.
b) μ is the molecular dipole (Debye). In parenthesis, values obtained with Gaussian03.
c) In parenthesis, volumes and populations of all-electron systems optimized at the B3LYP/aug-cc-pVDZ level of theory.

Molecular systems	Volume	Population [a]	$\|\mathbf{M}_1\|$	μ (D)[c]
V(C, C) basin:				
C_2H_6	18.4	1.84 (1.81) [a]	0.079	0.00(0.00)
C_2H_4	43.9	1.65 (1.69 [a]	0.620	0.00(0.00)
C_2H_2	167.5	5.08 (5.14) [a]	Circular[e]	0.00(0.00)
C_6H_6	53.5	2.77 (2.76) [a]	0.143	0.00(0.00)
NH_3-BH_3				5.35(5.379)
V(N, H)	54.9	1.97	1.06	
V(B, N)	18.8	1.91	0.483	
H_2O				
monomer				1.92 (1.92)
V(O)	49.0	2.29 (2.29) [a]	0.78(0.78) [a]	
dimer				2.43 (2.45)
$V(O_{acceptor})$	50.8	2.37(2.36) [a]	0.81(0.80) [a]	
$V(O_{acceptor})$	39.3	2.16(2.17) [a]	0.76(0.75) [a]	
$V(O_{donor})$	52.2	2.29(2.29) [a]	0.77(0.77) [a]	
H_2O_2				1.76(1.77)
V(O, O)	1.2	0.30(0.31)[b]	0.06	
C-O compounds				
V(C, O) - V(O) basins				
CH_3OH	6.9 - 46.5	1.25 - 2.37	0.07 - 0.85	1.61(1.61)
H_2CO	35.9 - 54.9	2.40 - 2.54	0.20 - 0.91	2.28(2.28)
$CO(^1\Sigma^+)$	37.4 – 300.4	3.04 - 4.21	0.21 - 3.00	0.14 (0.13)
CO_3^{2-}	13.8 - 54.2	2.01 - 2.54	0.12 - 1.01	0.00(0.00)
Glycine				
NH_2CH_2COOH				5.48(5.52)
V(C, O)	10.0/29.5	1.62/2.34	0.07/0.26	
V(O)	27.7	2.36	0.26	
V(N)	51.9	2.08	0.69	
$NH_3^+CH_2COO^-$				14.29(14.36)
V(C,O)	18.0	1.99	0.17	
V(O)	65.0	2.78	1.01	
$NH_3^+CH_2COO^-$ (50 H_2O)[g]				
$V(C,O_1)$	19.8	2.02	0.18	
$V(C,O_2)$	20.6	1.96	0.16	
$V(O_1)$	59.8	2.80	1.02	
$V(O_2)$	56.0	2.70	1.01	
glycine → H_2O[h]		0.12[g]		
$Cr(CH_3)_6$				0.07(0.08)
C(Cr)	28.3	21.15(21.11)[d]	0.67	
V(Cr)	0.82	2.10	0.08	
V(Cr, C)	25.4	1.25(1.26)[d]	0.10	
Open shell				
$CN(^2\Sigma^+)$				1.32(1.32)
V(N)	105.0	3.29	2.57	
V(C, N)	53.5	3.46	Circular[e]	
O_2 ($^3\Sigma^-_g$)				0.00(0.00)
V(O)	106.9	5.25 (5.2) [a]	Circular[e]	
V(O, O)	4.3	1.20 (1.30) [a]	0.01	
Methane Gas hydrate[f]				
CH_4				0.00
C(C)	0.80	2.09	0.00	
V(C, H)	94.0	1.97	1.00	
CH_4 in $(H_2O)_{20}$				0.00
C(C)	1.60	2.10	0.08	
V(C, H)	99.0	2.01	1.05	

Table 4. ELF population analysis from a cubic grid. Systems optimized at the B3LYP/cc-pVTZ level of theory except for the methane hydrate structure where the MPWB1K/cc-pVTZ level was employed. Populations are given in electrons and $|\mathbf{M}_1|$ are in a.u. The volumes (a.u.) are calculated with a threshold of 0.001 a.u. on the density. a) in parenthesis, the values were obtained with TopMod.

b) in parenthesis, the V(O, O) population was obtained with TopMod

c) μ is the molecular dipole given in Debye. The values given parenthesis are obtained with Gaussian03.

d) in parenthesis, values extracted from the reference: *Inorg. Chem.* **2004**, 43, 3248-3256.

e) the basin attractor is circular (not punctual) and thus, the $\mathbf{M}_1$ quantity cannot be properly calculated (see reference *J. Comput. Chem.* **2008**, 29, 9, 1440-1449)

f) methane hydrate optimized at the MPWB1K/6-31+G(d,p) level of theory. The ELF analysis was performed using a medium grid step used of 0.09 a.u. g) Basins numbering with respect to the numbering given in Figure 3. h) Charge transfer from oxygen lone pairs to surrounded water molecules calculated as follows : $[\Sigma\ pop.V(O)]_{zwitterion+50H2O} - [\Sigma\ pop.\ V(O)]_{zwitterion}$

molecule or crystal	Volume	N (electrons)	$\lvert\mathbf{M}_1\rvert$
C_2H_4			
C	335.0	4.04	0.67
V(C, C)	47.0(43.9)[a]	1.84(1.64)[a]	0.60(0.62)[a]
V(C, H)	75.0	1.98	0.22
C_6H_6			
C	81.2	4.01	0.35
V(C, C)	55.0(53.5)[a]	2.83(2.77)[a]	0.15(0.14)[a]
Diamond crystal			
V(C, C)	18.6(20.4)[b]	1.92(1.98)[b]	0.15
ZnS crystal			
C(Zn)	6.6	7.2	0.38
V(Zn)	300.2	4.65	15.9
C(S)	6.5	0.93	1.09
V(S)	44.0	2.04	1.20

Table 5: QTAIM and ELF populations obtained from CHGCAR and ELFCAR output files.
a) Values in parenthesis are obtained with TopMod at the B3LYP/cc-pVTZ level of theory.
b) values in parenthesis are extracted from reference*: J. Chem. Theory Comput*., 2009, 5 (1), 164-173

	ELF			L(r)			MEP		
	Attractor[a]	Volume	Pop.	Attractor[a]	Volume	Pop.	Attractor[1]	Volume	Pop.
Oxygen core	0.0	0.25	2.13	0.00	0.02(0.01)[b]	1.14(0.76)[b]	0.0	26.6	7.58
Hydrogen core	-	-	-	0.96	3.22(3.20)[b]	0.57(0.57)[b]	-	-	-
O-H bond	0.96	44.8	1.65	0.38	0.99(0.80)[b]	0.56(0.52)[b]	0.96	43.5	0.88
Oxygen lone pair	0.59	49.0	2.29	0.33	2.61(2.60)[b]	1.70(1.83)[b]	1.23	37.1	0.33

Table 6: Compared ELF, $L(r)=-\nabla^2(r)$ and MEP populations analysis.
Populations are given in electrons and volumes (a.u.) are calculated with a threshold of 0.001 a.u. on the density. The grid step used is 0.06 a.u.
a) Distance (Å) of the basin attractor from the oxygen nuclear center.
b) Values in parenthesis are extracted from the reference: Nathaniel O. J. Malcolm and Paul L. A. Popelier. *Faraday Discuss.*, 2003, 124, 353-363

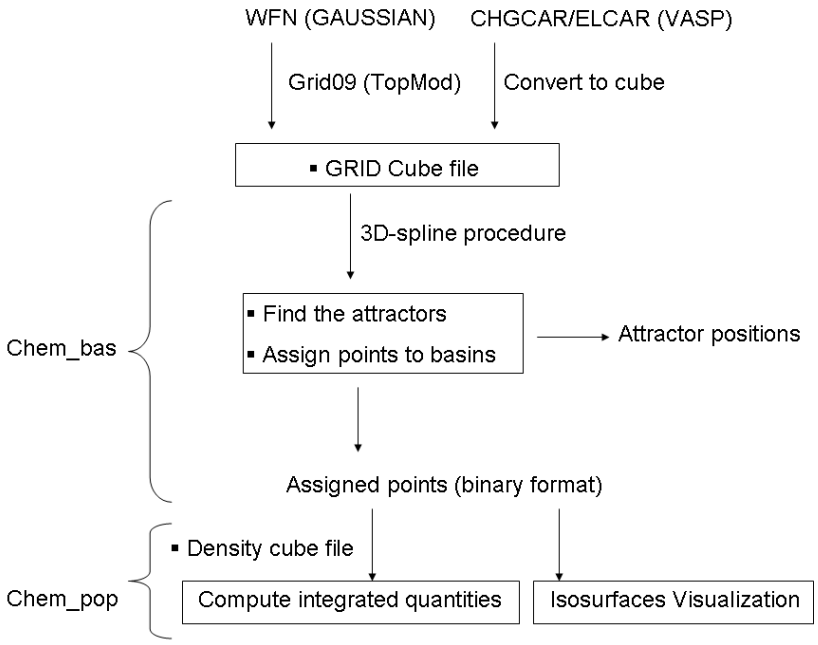

Figure 1: Flow diagram of the organization of the modules in TopChem.

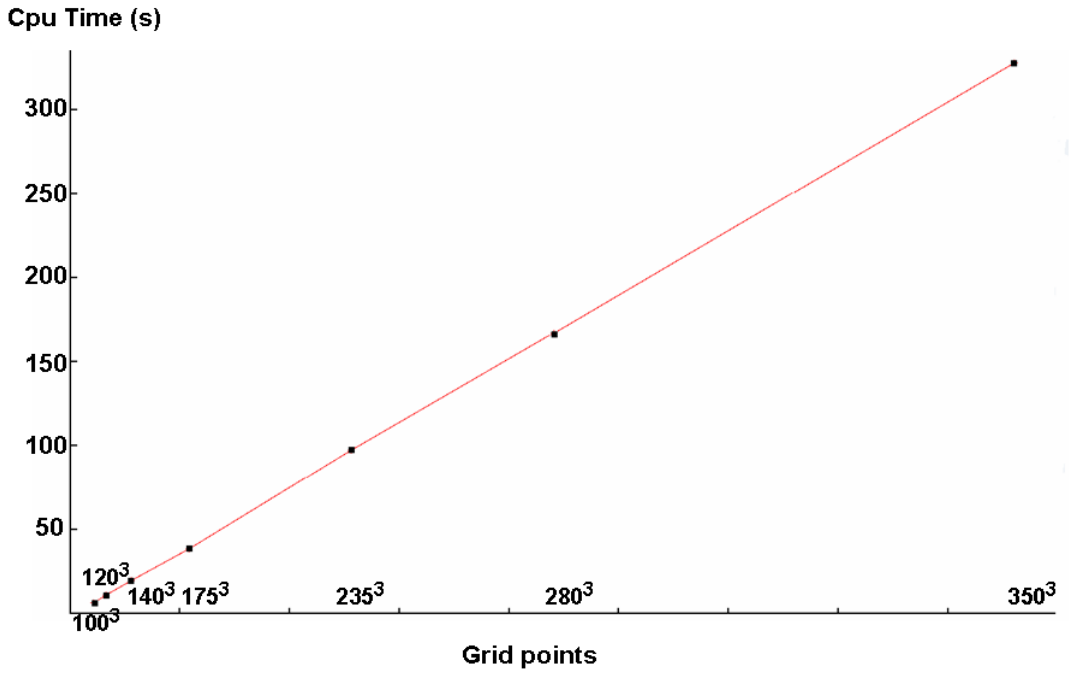

Figure 2: Timing data for the ELF analysis of the CH_3OH molecule. The cpu time scales linearly with the number of grid points.

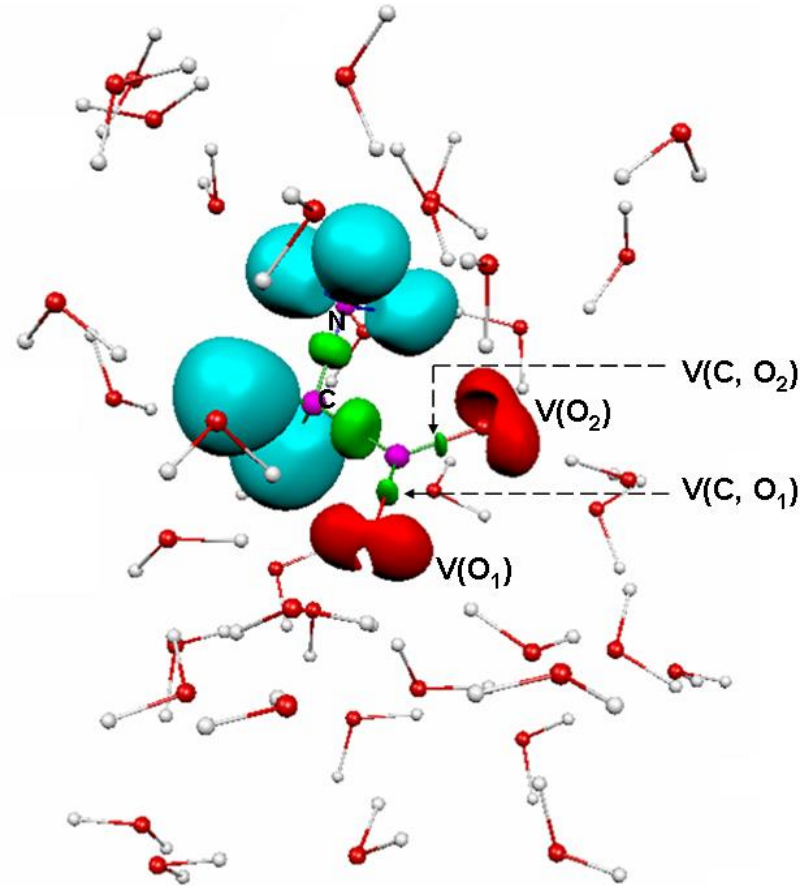

Figure 3: ELF localization domains (ELF = 0.85) of the zwitterionic glycine embedded in 50 water molecules and PCM continuum. Optimized structural parameters were obtained with a combined approach B3LYP/PM6 using a cc-pVTZ basis set for the high level theory. Color code: magenta: core C(C), C(O) and C(N) basins; green: V(C, O), V(C, N) and V(C, C) bonding basins; light blue; V(C, H) and V(N, H) protonated basins; red: V(O) oxygen lone pair basins**.** The basins of water molecules were omitted for clarity.

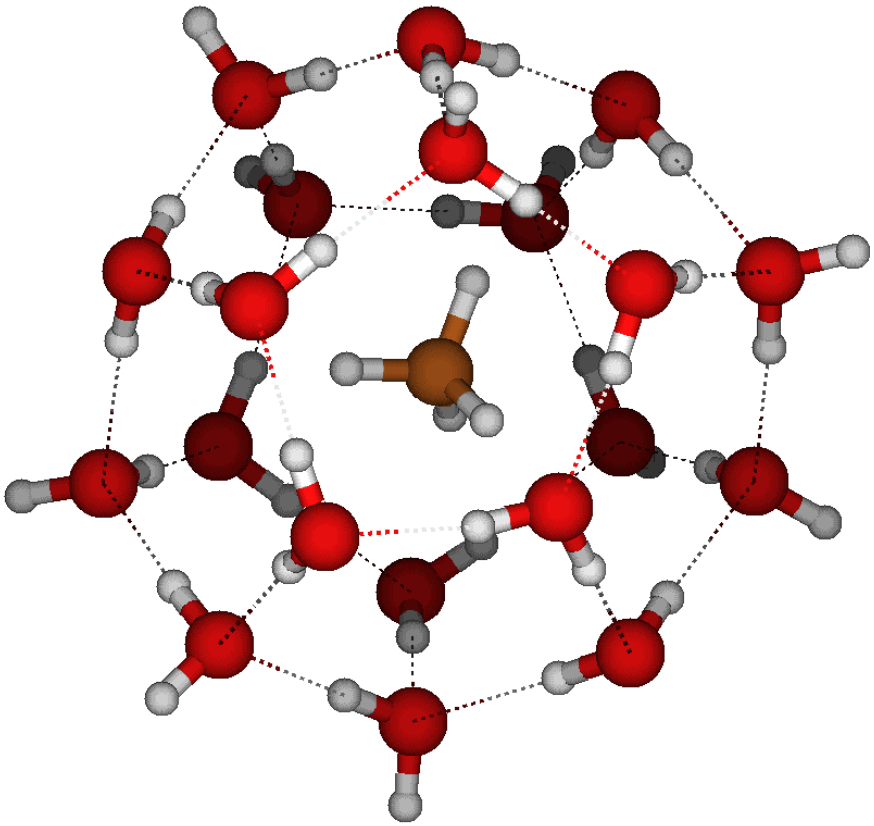

Figure 4: Methane gas hydrate structure optimized at the MPWB1K/6-31+G(d,p) level of theory.

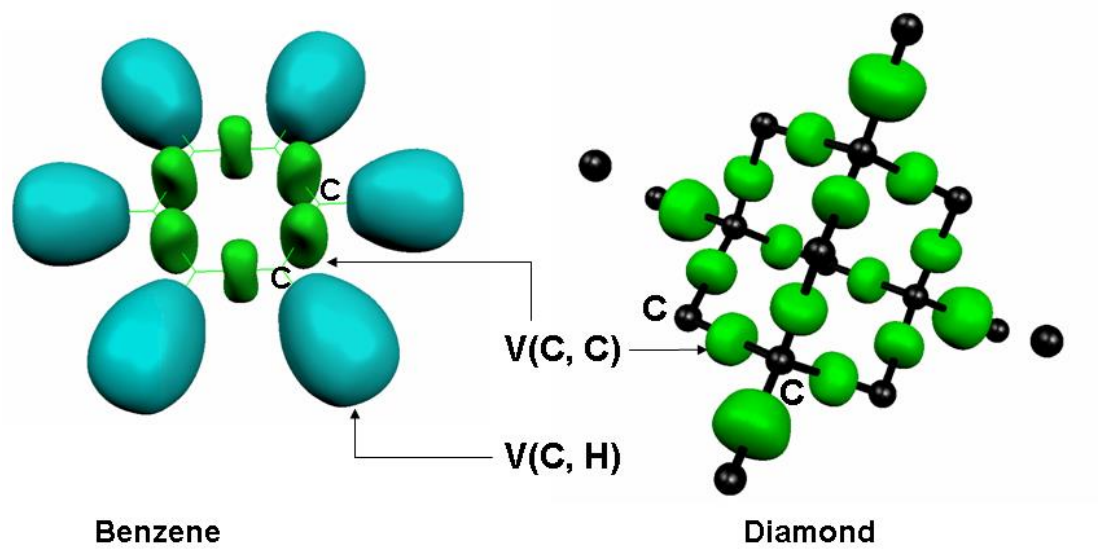

Figure 5: ELF localization domains (ELF = 0.85) of C_6H_6 and diamond crystal. Color code: green: V(C, C) bonding basin, light blue : V(C, H) protonated basins.

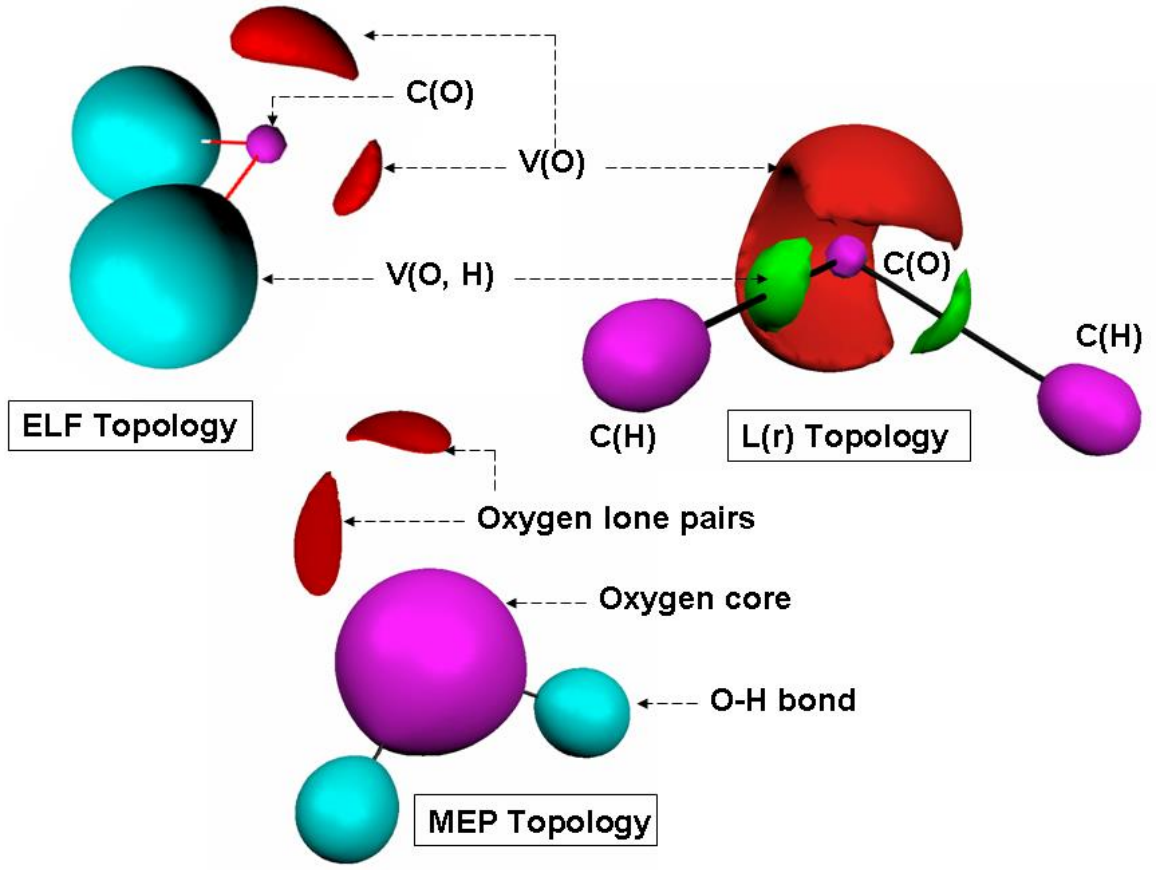

Figure 6: Localization domains of ELF, L(r) and MEP (positive and negative were merged). Color code: magenta: core regions; green: O-H bonds, light blue: protonated O-H bonds and red: oxygen lone pairs.