Electron density analysis of large (molecular and periodic) systems: A parallel implementation

Electron Density Analysis of Large (Molecular and Periodic) Systems: A Parallel Implementation.

S. Casassa,[1,*] A. Erba,[1] J. Baima,[1] and R. Orlando[1]

[1] *Dipartimento di Chimica, Università di Torino and NIS, Nanostructured Interfaces and Surfaces, Centre of Excellence, Via Giuria 5, 10125 Torino, Italy*

(Dated: July 14, 2015)

A parallel implementation is presented of a series of algorithms for the evaluation of several one-electron properties of large molecular and periodic (of any dimensionality) systems. The electron charge and momentum densities of the system, the electrostatic potential, X-ray structure factors, directional Compton profiles can be effectively evaluated at low computational cost along with a full topological analysis of the electron charge density of the system according to Bader's quantum theory of atoms in molecules. The speedup of the parallelization of the different algorithms is presented. The search of all symmetry-irreducible critical points of the electron charge density of the crystallized crambin protein and the evaluation of all the corresponding bond paths, for instance, would require about 32 days if run in serial mode and reduces to less than 2 days when run in parallel mode over 32 processors.

I. INTRODUCTION

Through the *ab initio* quantum-chemical solution of the static Schrödinger equation $\hat{H}|\Psi\rangle = E|\Psi\rangle$ of the ground-state of a molecular or periodic system, much attention is commonly devoted to the total electronic energy E and to its derivatives as they allow to determine a variety of properties of interest (equilibrium structures, binding energies, relative stabilities, vibrational spectra, thermodynamic functions, elastic constants, etc.). Nonetheless, a wealth of chemical information on the system can be fruitfully extracted from the analysis of the corresponding wave-function $|\Psi\rangle$. Being a rather complex function of N space-spin coordinates (with N number of electrons), $|\Psi\rangle$ is commonly replaced by n-particle density matrices Γ_n (functions of just $2n$ space-spin variables), which allow for the evaluation of the expectation value of any n-particle operator. In particular, from the one-electron density matrix $\Gamma_1(\mathbf{x}, \mathbf{x}')$ any one-electron property of the system can be effectively computed (electron charge density, ECD $\rho(\mathbf{r})$, X-ray structure factors, electrostatic potential, electron momentum density, EMD $\pi(\mathbf{p})$, Compton profiles, etc.).[1]

Among one-electron properties, the ECD is by far the most studied as it embodies rich information on the structure and chemical nature of the bonding of the system. From a fundamental point of view, the density functional theory (DFT) has given the ECD an even more relevant role in quantum chemistry as the energy of the system has been demonstrated to be a functional of it: $E[\rho]$.[2,3] The ECD $\rho(\mathbf{r})$ is a relatively simple function of just three position space coordinates, which exhibits the point- or group-symmetry of the system. Several schemes have been proposed in the last decades to extract from $\rho(\mathbf{r})$ as much chemical information as possible, which involve its visualization (profiles, 2D maps, 3D graphical representations), partition according to different schemes, topological analysis, etc.[4–15] In this respect, one of the most powerful and popular techniques is represented by Bader's quantum theory of atoms in molecules (QTAIM), which relies on the topological analysis of $\rho(\mathbf{r})$.[16]

If a routine analysis of the wave-function $|\Psi\rangle$ of a small system is a non particularly demanding task from a computational point of view (when compared to the convergence of the self-consistent field, SCF, procedure or to the analytical evaluation of energy gradients), this is clearly no more the case either when rather sophisticated techniques are adopted (such as QTAIM) or when large systems are studied. In this respect, one of the challenges state-of-the-art *ab initio* quantum chemistry is facing is indeed that of extending its applicability range to realistic systems of large dimensions for which the analysis of $|\Psi\rangle$ can easily take several days of computation if run in serial mode.

In this paper, we present an efficient parallel implementation, in the CRYSTAL program, of a series of density matrix-based algorithms for computing, at the *ab initio* level of theory, a variety of one-electron properties of large systems. The present implementation takes advantage of the several significant improvements recently made in CRYSTAL in terms of increased parallel and massive-parallel scalability, reduced use of memory per node and increased exploitation of symmetry at all steps of the calculation, which recently allowed the program to be run in parallel mode over 32'000 CPUs and to study systems containing up to about 14'000 atoms and 200'000 basis functions.[17–19] Specific features of the current implementation are: i) possibility of studying systems of any dimensionality within the same formal and numerical framework (from 0D molecules, to 1D polymers, nanotubes, helices and nano-rods, to 2D slabs and 3D crystals); ii) efficient use of several DFT functionals, belonging to four rungs of the well-known "Jacob's ladder"[20] (local density approximation, LDA, generalized gradient approximation, GGA, global or range-separated hybrids and meta-GGA); iii) full exploitation of any residual symmetry; iv) parallelization of all algorithms related to the evaluation of $\rho(\mathbf{r})$, its gradient and Laplacian, of

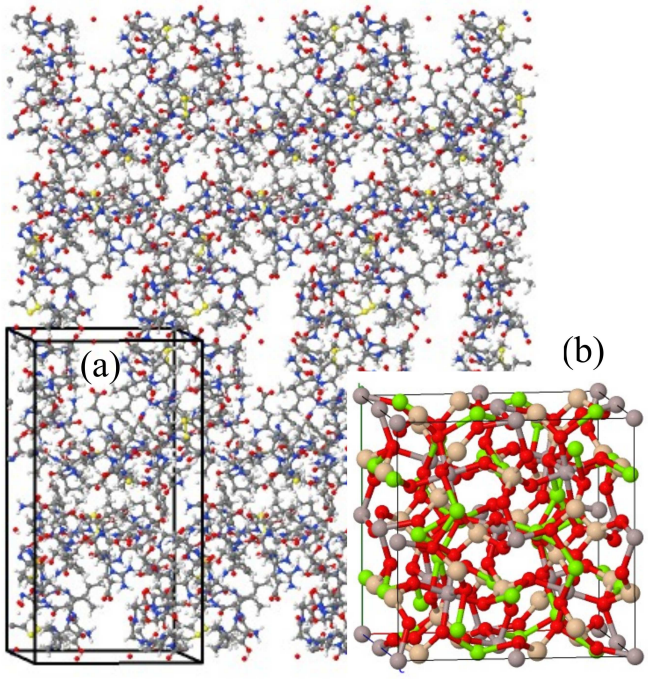

FIG. 1: (color online) Graphical representation of the crystal structure of (a) the crambin protein and (b) the pyrope silicate garnet. The crambin unit cell contains two crambin molecules.

the X-ray structure factors, of Bader's topological analysis (as generalized to periodic systems by C. Gatti's TOPOND package,[21,22] which has recently been merged into the CRYSTAL program), of directional Compton profiles, of the electrostatic potential and its derivatives, of the electronic band structure and density-of-states. The CRYSTAL program adopts an atom-centered basis set of Gaussian-type functions (GTF); all density matrix-based algorithms have been parallelized on the number of orbital shell-shell pairs, which guarantees a good load balance among processors and thus a satisfactory speedup for most systems.

We consider two rather different 3D systems in order to discuss the scalability of the present parallel implementation: the pyrope silicate garnet and the crystalline form of the crambin protein. Pyrope, $Mg_3Al_2(SiO_4)_3$, is the most abundant of silicate garnets, which are among the main constituents of the Earth's lower crust, upper mantle and transition zone. This family of crystals is characterized by a cubic structure with space group $Ia\overline{3}d$ and formula $X_3Y_2(SiO_4)_3$, where the X site hosts divalent cations such as Ca^{2+}, Mg^{2+}, Fe^{2+} and Mn^{2+} and the Y site is occupied by trivalent cations such as Al^{3+}, Fe^{3+} and Cr^{3+}.[23,24] The primitive cell contains four formula units (80 atoms, for a total of 1488 atomic orbitals with the present basis set) and the structure consists of a network of corner-sharing SiO_4 tetrahedra and YO_6 octahedra. Crambin is a 46-residues protein, belonging to the family of thionins, which can be extracted from the Abyssinian cabbage and crystallizes in a well-resolved structure with two proteins per unit cell,[25] for a total of 1284 atoms per cell (corresponding to 12'354 atomic orbitals with the present basis set). The crambin molecule has recently been used as a test-case to investigate the convergence of atom-atom electrostatic interaction energy from the computation of high-rank topological atomic multipole moments.[26] The atomic structure of the two crystals is graphically represented in Figure 1.

The structure of the paper is as follows: Section II is devoted to the formal presentation of the one-electron properties whose algorithms have been parallelized in the present implementation; in Section III we will discuss the speedup of the parallel implementation along with some specific features of the present implementation; in Section IV conclusions are drawn.

II. COMPUTATIONAL METHODOLOGY AND SETUP

A. One-electron Properties

The present parallelized implementation of the wave-function analysis of molecules and periodic systems of different dimensionality (1D, 2D, 3D) has been made into the CRYSTAL14 program, which adopts a basis set of atom-centered GTFs. Within a linear-combination-of-atomic-orbitals scheme, crystalline orbitals can be expressed as:

$$\psi^{j,\kappa}(\mathbf{r}) = \sum_{\mu} a_{\mu}^{j,\kappa} \left[\sum_{\mathbf{g}} e^{\imath\,\kappa\cdot\mathbf{g}}\, \chi_{\mu}^{\mathbf{g}}(\mathbf{r}) \right] , \qquad (1)$$

where the band index j runs from 1 to $N/2$ (for a closed-shell system), $\chi_{\mu}^{\mathbf{g}}$ is an atomic orbital centered on a cell identified by the $\mathbf{g}$ lattice vector and the wave-vector κ is one of the L vectors in the first Brillouin Zone of reciprocal space, which correspond to the number of cells in the cyclic crystal determined by the Born von Kármán periodic boundary conditions. The one-electron density matrix of the system can then be expressed as follows, in a matrix representation on the basis of the atomic orbitals:

$$P_{\mu\nu}^{\mathbf{g}} \;=\; 2 \sum_{j=1}^{N/2} \sum_{\kappa} e^{-\imath\,\kappa\cdot\mathbf{g}} \left[a_{\mu}^{j,\kappa} \left(a_{\nu}^{j,\kappa} \right)^{*} \right] . \qquad (2)$$

The ECD $\rho(\mathbf{r})$ of the system is evaluated at each position $\mathbf{r}$ of direct space as:[27]

$$\rho(\mathbf{r}) = \sum_{\mathbf{g},\mathbf{l}} \sum_{\mu,\nu} P_{\mu\nu}^{\mathbf{g}-\mathbf{l}} \chi_{\mu}^{\mathbf{g}}(\mathbf{r}) \chi_{\nu}^{\mathbf{l}}(\mathbf{r}) \, . \qquad (3)$$

Depending on the selected set of $\mathbf{r}$ points where the above expression is evaluated, 2D maps or 3D plots of the ECD

can easily be obtained. X-ray structure factors $\{F_{hkl}\}$ are given by the Fourier transform of the ECD as:[28]

$$F_{hkl} = \int \rho(\mathbf{r}) e^{\imath\, \kappa\cdot\mathbf{r}} \mathrm{d}\mathbf{r} \ , \tag{4}$$

with $\kappa = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3$ being one of the lattice points of reciprocal space ($\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ being the three fundamental reciprocal lattice vectors). The EMD $\pi(\mathbf{p})$ of the system is given by:[29]

$$\pi(\mathbf{p}) = \sum_{\mu,\nu}\sum_{\mathbf{g}} P^{\mathbf{g}}_{\mu\nu} e^{\imath\mathbf{p}\cdot(\mathbf{r}_\nu - \mathbf{r}_\mu - \mathbf{g})}\, \chi_\mu(\mathbf{p})\, [\chi_\nu(\mathbf{p})]^* \ , \tag{5}$$

where $\chi_\mu(\mathbf{p})$ is the Fourier transform of the atomic orbital $\chi_\mu(\mathbf{r})$ centered on the origin of the spatial coordinates and $\mathbf{r}_\mu$ is the position the atomic orbital χ_μ is centered on within the reference cell. For the analysis and manipulation of the EMD, reference can also be made to the auto-correlation function, $B(\mathbf{r})$, first introduced by Pattison *et al.* in 1977.[30] The auto-correlation function is defined as the 3D Fourier transform of the EMD or as the auto-correlation integral of the position density matrix and can be given the following definition in the basis of the atomic orbitals:

$$\begin{aligned} B(\mathbf{r}) &= \sum_{\mu,\nu}\sum_{\mathbf{g}} P^{\mathbf{g}}_{\mu\nu}\, S^{\mathbf{g}}_{\mu\nu}(\mathbf{r}) \ , \text{ with:} \\ S^{\mathbf{g}}_{\mu\nu}(\mathbf{r}) &= \int \chi^{\mathbf{0}}_\mu(\mathbf{r}')\, \chi^{\mathbf{g}}_\nu(\mathbf{r}+\mathbf{r}')\, d\mathbf{r}' \ . \end{aligned} \tag{6}$$

Directional Compton profiles can be efficiently computed from the $B(\mathbf{r})$ function.[31–33] Also the electrostatic potential $V(\mathbf{r})$ can be evaluated from the knowledge of the one-electron density matrix. While the nuclear term is evaluated via the usual Ewald scheme for the calculation of the periodic lattice interactions, the electronic term is given by:

$$V_{\mathrm{ele}}(\mathbf{r}) = \sum_{\mu,\nu}\sum_{\mathbf{g}} P^{\mathbf{g}}_{\mu\nu} \left[\sum_{\mathbf{l}} \int \frac{\chi_\mu(\mathbf{r}')\chi_\nu(\mathbf{r}'-\mathbf{g})}{|\mathbf{r}' - (\mathbf{r}-\mathbf{l})|} \mathrm{d}\mathbf{r}' \right] . \tag{7}$$

All the algorithms for the evaluation of the above mentioned density matrix-related one-electron properties have been parallelized on $(\mu\nu)$ orbital shell pairs.

B. Topological ECD Analysis

Bader's QTAIM probably represents the most popular and complete theoretical framework for the topological analysis of the ECD of a molecular system.[16] Its generalization to periodic systems, as implemented since 1998 by C. Gatti in the TOPOND package,[22,34] has recently been fully merged into the CRYSTAL program.[17] A fundamental aspect of the topological analysis of the ECD is the determination of its critical points (CP), *i.e.* those points in space where $\nabla\rho(\mathbf{r}) = 0$. Any CP can be classified in terms of the eigenvalues λ_1, λ_2 and λ_3 of the Hessian matrix of the ECD second-derivatives evaluated at the CP; accordingly, each CP is labeled by a combined index (r, s) where r is the number of non-zero eigenvalues λ_i ($r = 3$ for all CPs of stable 3D molecular or solid structures) and s is the difference between the number of positive and negative eigenvalues. The value of s can be 3 for a minimum (all three eigenvalues are positive), -3 for a maximum (all three eigenvalues are negative) and +1 or -1 for saddle points in one or two dimensions. Among others, critical points of (3,-1) type (*i.e.* bond critical points) are of particular chemical interest. By default, CPs are searched using a standard Newton-Raphson scheme;[22] a more sophisticated *eigenvector following* approach can also be used.[35,36] Several quantities can be evaluated at CPs, such as the Laplacian $\nabla^2\rho(\mathbf{r})$, the potential energy density $V(\mathbf{r})$, the positive definite kinetic energy density $G(\mathbf{r})$, and the electronic energy density $H(\mathbf{r}) = V(\mathbf{r}) + G(\mathbf{r})$, which are related to the ECD Laplacian through the local virial theorem:[34]

$$\frac{1}{4}\nabla^2\rho(\mathbf{r}) = V(\mathbf{r}) + 2G(\mathbf{r}) = H(\mathbf{r}) + G(\mathbf{r}) \ . \tag{8}$$

The $|V(\mathbf{r})|/G(\mathbf{r})$ ratio and the bond degree (BD) index $H(\mathbf{r})/\rho(\mathbf{r})$ also prove useful in the chemical description of different bonds, as the nature of the interactions can be rationalized as follows:[37] i) covalent bonds exhibit negative Laplacian, negative $H(\mathbf{r})$ and a $|V(\mathbf{r})|/G(\mathbf{r})$ ratio larger than 2; ii) transit bonds are associated with a positive Laplacian, an almost zero value of BD and a value of $|V(\mathbf{r})|/G(\mathbf{r})$ between 1 and 2; (iii) ionic bonds, hydrogen-bonds and van der Waals interactions show positive Laplacian and $H(\mathbf{r})$ and a $|V(\mathbf{r})|/G(\mathbf{r})$ ratio lower than 1. Integration over atomic basins gives further information such as atomic volumes, Bader's atomic charges and the partition of the energy density in atomic contributions.

C. Computational Setup

All calculations are performed with the CRYSTAL14 program.[17] All-electron atom-centered GTF basis sets are adopted. For pyrope, oxygen atoms are described by a (8*s*)-(411*sp*)-(1*d*) contraction of primitive GTFs, silicon by a (8*s*)-(6311*sp*)-(1*d*) one, aluminum by a (8*s*)-(611*sp*)-(1*d*) one and magnesium by a (8*s*)-(511*sp*)-(1*d*) one. A 6-21G(*d*) basis set is used for the crambin crystal. A Linux cluster of Intel-Xeon processors, working at 2.13 GHz, with Ethernet connection is used for all calculations. The machine has 8 cores per node and 2 GB of memory per core.

III. RESULTS AND DISCUSSION

As recalled in the introduction, the most interesting one-electron property of a system is by far the ECD $\rho(\mathbf{r})$, out of which a wealth of chemical information can be

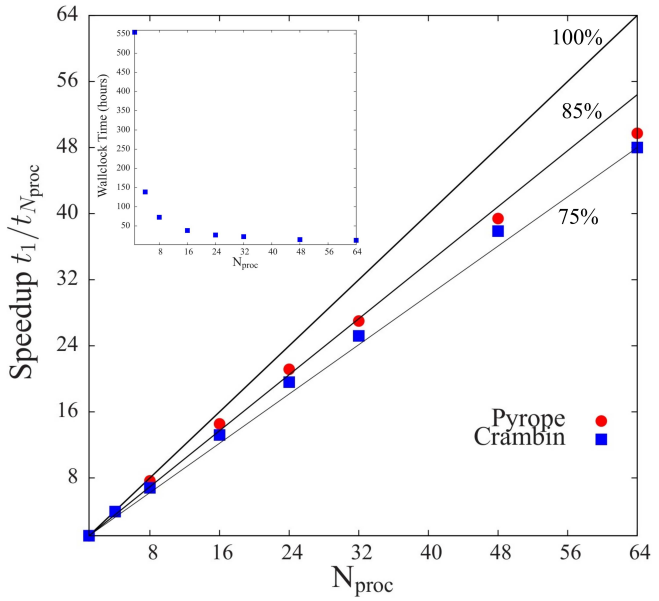

FIG. 2: (color online) Speedup of the calculation of the ECD $\rho(\mathbf{r})$ on a 500×500×500 3D grid of points of pyrope (red circles) and crambin (blue squares) as a function of the number of processors used N_{proc}. The three black lines of decreasing thickness mark the 100%, 85% and 75% efficiency of the parallelism, respectively. The inset shows the reduction of the wall-clock time of the calculation in the case of crambin as a function of N_{proc}.

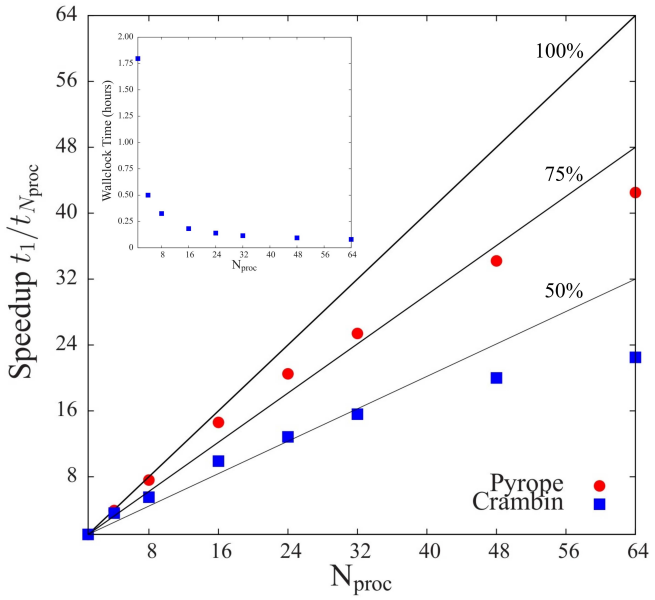

FIG. 3: (color online) Speedup of the calculation of the electrostatic potential on 3D grid of 250'000 points of pyrope (red circles) and crambin (blue squares) as a function of the number of processors used N_{proc}. The three black lines of decreasing thickness mark the 100%, 75% and 50% efficiency of the parallelism, respectively. The inset shows the reduction of the wall-clock time of the calculation in the case of crambin as a function of N_{proc}.

extracted. The efficiency of the current parallel implementation of its calculation according to equation (3) is shown in Figure 2 where we report the speedup of the calculation as a function of the number of processors N_{proc} used (up to 64). The speedup is defined as the $t_1/t_{N_{\mathrm{proc}}}$ ratio between the wall-clock time needed to run the calculation in serial mode t_1 (*i.e.* on one processor) and that needed to run it in parallel mode on N_{proc} processors $t_{N_{\mathrm{proc}}}$. The speedup is given for both the pyrope silicate garnet (red circles) and for the crambin crystallized protein (blue squares) and corresponds to the evaluation of $\rho(\mathbf{r})$ on a 500×500×500 3D grid of points. A high efficiency of the parallelization (greater than 75%) is observed for both systems up to 64 processors, pyrope showing a slightly larger speedup than crambin. Up to 16 processors, the efficiency is above 85% for both systems. The inset of the figure reports the absolute wall-clock time required by such a calculation in the case of crambin (where such a dense grid of points corresponds to a spatial resolution of 0.08×0.03×0.04 Å) as a function of the number of cores used. It is seen that about 550 hours (*i.e.* about 23 days!) would be required to compute such property on serial mode, whereas it only takes 10 hours for a calculation run in parallel mode on 64 processors.

In general, the evaluation of the electrostatic potential is not a prohibitive computational task; however, for large systems it might become demanding. In Figure 3, we report the speedup of the evaluation of the electrostatic potential (whose electronic contribution is given by equation 7) at a 3D grid of 250'000 points for pyrope and crambin. In this case, it is seen that the speedup of the calculation significantly reduces when passing from pyrope to crambin, as there are small parts of the algorithm which have not been parallelized. Anyhow, the efficiency of the scaling is seen to be always larger than 75% and 50% for pyrope and crambin, respectively up to 32 cores. Above 32 cores the speedup flattens, particularly so in the case of crambin. The inset of the figure reports the absolute wall-clock time of such a calculation for crambin; it is seen that it would require about two hours if run in serial mode, reducing to about 5 minutes if run in parallel mode on 64 processors. The evaluation of the ECD of a system on a 3D grid of points might be performed when one wants to obtain 3D plots of $\rho(\mathbf{r})$. A standard option of this kind of graphical representations is that of coloring the ECD iso-density surfaces according to the value of the electrostatic potential at each position in space, as it constitutes a useful tool for identifying possible active sites in complex structures. By combining the two algorithms discussed above, these representations can be achieved at low computational cost and short wall-clock time even for large molecular or periodic systems. Figure 4 shows such a graphical representation for the cell of the crambin protein crystal.

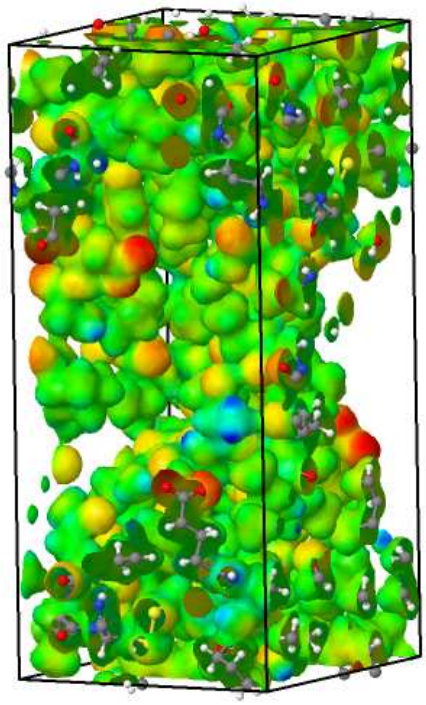

FIG. 4: (color online) 3D graphical representation of an isodensity ECD surface of the crambin crystal, colored according to the value of the electrostatic potential at each point of space.

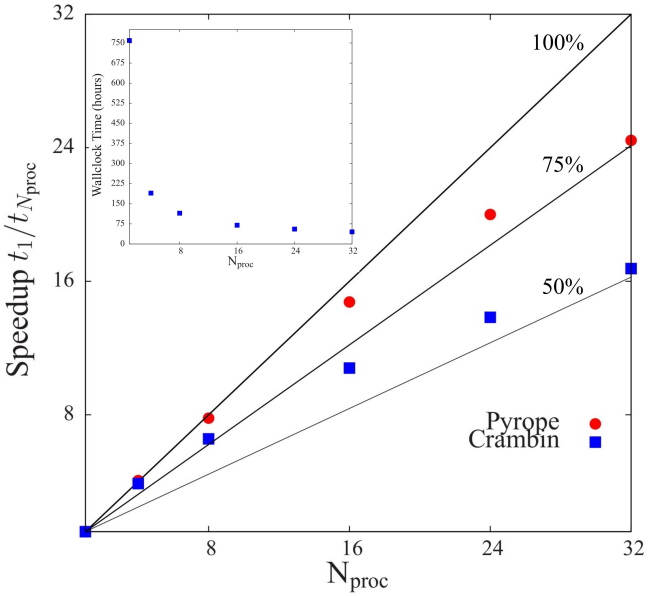

FIG. 5: (color online) Speedup of the determination of ECD critical points and *bond paths* of pyrope (red circles) and crambin (blue squares) as a function of the number of processors used $N_{\rm proc}$. The three black lines of decreasing thickness mark the 100%, 75% and 50% efficiency of the parallelism, respectively. The inset shows the reduction of the wall-clock time of the calculation in the case of crambin as a function of $N_{\rm proc}$.

A. Topological Analysis

As recalled in Section II B, the critical points of the ECD $\rho(\mathbf{r})$ can be classified in different types depending on the number of negative eigenvalues of the Hessian matrix computed at the critical point. Bond critical points, of (3,-1) type, are of particular interest due to the fact that several indicators computed at those points can give valuable information on the particular chemical nature of the interatomic interactions. Generally, one wants to determine CPs of just a chemically meaningful portion of the system. However, such determination might become a demanding computational task if a large portion of the system is considered and if *bond paths* are to be computed. Let us consider the whole cell of both pyrope and crambin: pyrope has 18 symmetry-irreducible critical points per cell while for crambin they are as many as 2052. A serial calculation for the search of the CPs and for the evaluation of the *bond paths* would require about 40 minutes for pyrope and 760 hours for crambin. Figure 5 reports the speedup of this calculation as a function of the number of processors used, up to 32. It is seen that up to 8 processors the scaling of pyrope is almost ideal, with an efficiency close to 100%, and that of crambin is above 75%; from 8 to 32 processors the efficiency of the parallelization decreases but still remains larger than 75% for pyrope and 50% for crambin. This speedup allows to determine the critical points and the *bond paths* of the crambin crystallized protein in a bit less than 2 days when running in parallel on 32 processors with respect to the 32 days that would be required in serial mode. The reduction of the wall-clock time on such a calculation is documented in the inset of the figure.

As introduced in Section II B, several quantities of chemical interest (atomic partition of the kinetic and potential energy density, Bader's atomic charges, atomic volumes, etc.) can be extracted from the analysis of the ECD by performing an integration over atomic basins. This part of the ECD topological analysis is by far the most computationally expensive; if the symmetry-irreducible portion of the system is large, it is generally convenient to select a chemically-interesting fragment of the system on which to perform such an analysis. Pyrope has a relatively small asymmetric unit so that the integration has been performed over all of it, whereas for crambin a tyrosine fragment has been selected that is known to constitute one of the active sites for water absorption in the crystalline structure. The default value of 0.001 a.u. is used for the tolerance that governs the determination of zero flux surfaces in the TOPOND package, which generally ensures very high accuracy; it may be noticed that less strict values of this tolerance (approximately in the range from 0.002 to 0.004 a.u.) could be safely used, which would reduce the required CPU time. In Figure 6 we report the speedup of such a calculation up to 32 processors; the selected tyrosine fragment of crambin is graphically represented in the inset of the

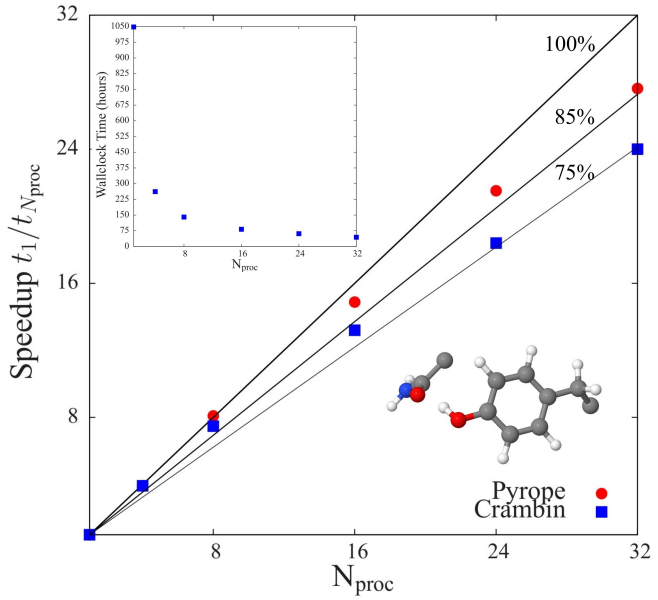

FIG. 6: (color online) Speedup of the calculation of integrated ECD-related properties of pyrope (red circles) and of the tyrosine fragment (in the inset) of crambin (blue squares) as a function of the number of processors used $N_{\rm proc}$. The three black lines of decreasing thickness mark the 100%, 85% and 75% efficiency of the parallelism, respectively. The inset shows the reduction of the wall-clock time of the calculation in the case of crambin as a function of $N_{\rm proc}$.

figure. It is seen that the efficiency of the speedup is above 85% for both systems almost up to 16 processors; up to 32 processors it remains larger than 85% for pyrope and it reduces to 75% for crambin. This speedup proves extremely useful for practical purposes, given the computationally demanding character of these calculations. As in the previous figures, also in this case the inset shows the reduction in wall-clock time of the calculation in the case of crambin. It is seen that such a characterization would require almost 44 days (1048 hours) if run in serial mode on one processor; the calculation only takes 2 days if run in parallel mode over 32 processors or 1 day if run over 128 processors (not shown in the inset). In the case of pyrope, the calculation takes about 32 hours on 1 processor and just 1.1 hours in parallel on 32 processors.

As a final consideration on the efficiency of the current parallelization, we may note that even if all speedups are rather satisfactory and effectively serve the practical purpose of making all of the present calculations feasible in a reasonable amount of time, none of them are actually ideal for large systems. This is due to the fact that some small parts of the algorithms (as the construction of the grid) have not been parallelized yet, which are almost negligible for small systems but deteriorate a bit the scaling for larger systems.

For practical purposes, a critical discussion of the relative performance of different functionals on the description of the chemical nature of various interactions does constitute a relevant quantum-chemical topic. In this respect, a further point of strength of the present scheme is given by the fact that it takes advantage of the progresses made in the last years in the implementation into the CRYSTAL program of several functionals of the DFT.[17] In particular, CRYSTAL is well-known for its efficient evaluation of the non-local Hartree-Fock exact exchange and for its inclusion in global or range-separated hybrid functionals. Four rungs of the so-called "Jacob's ladder" proposed by John Perdew to rationalize various DFT functionals in terms of increased accuracy[20] are currently implemented in the program: the local density approximation (LDA),[38] several formulations of the generalized-gradient approximation (GGA), such as the popular PBE or PW91 functionals,[39,40] global hybrids, such as B3LYP and PBE0,[41,42] and meta-GGA, such as the M06 one.[43]

TABLE I: Several properties (ECD ρ, its Laplacian $\nabla^2\rho$, the $|V|/G$ ratio and the bond degree H/ρ, all in atomic units) computed at two bond critical points (a covalent O−H, whose length is 0.99 Å, and a hydrogen O···H one, whose length is 1.83 Å) of the tyrosine fragment of crambin with different DFT functionals at fixed geometry (B3LYP optimized). For the O···H bond the strength of the hydrogen bond is estimated by its energy E (in kJ/mol). The fractional distance $x_{\rm CP}$ of the CP from the O atom, given with respect to the length of the bond, is also reported for both bonds.

	O−H					O···H					
	ρ	$\nabla^2\rho$	$\|V\|/G$	H/ρ	$x_{\rm CP}$	ρ	$\nabla^2\rho$	$\|V\|/G$	H/ρ	E	$x_{\rm CP}$
LDA	0.311	-1.96	9.66	-1.78	0.82	0.032	0.10	0.93	0.05	-28.2	0.66
PBE	0.320	-1.89	9.11	-1.69	0.81	0.031	0.10	0.89	0.08	-27.6	0.65
PW91	0.320	-1.89	9.11	-1.69	0.81	0.031	0.10	0.90	0.08	-27.6	0.65
B3LYP	0.320	-1.97	9.80	-1.74	0.81	0.031	0.11	0.88	0.09	-27.9	0.65
PBE0	0.320	-2.01	10.05	-1.77	0.82	0.030	0.11	0.88	0.10	-27.8	0.66
M06	0.310	-1.78	8.61	-1.65	0.81	0.033	0.11	0.92	0.06	-29.7	0.65

In Table III A we report a series of ECD-related properties as evaluated at two (3,-1) bond critical points of the tyrosine fragment of the crambin protein as computed with these 6 different functionals at fixed geometry (as optimized at B3LYP level). The ECD ρ, its Laplacian $\nabla^2\rho$, the $|V|/G$ ratio and the bond degree H/ρ are given for the covalent O−H bond and for the O$\cdots$H hydrogen bond; for the latter, its strength is also estimated by means of its energy $E = 1/2V$ (where V is evaluated at the corresponding critical point) according to the proposal by Espinosa *et al.*[44] The position of the two CPs is also reported in the table, expressed as the fractional distance $x_{\rm CP}$ from the O atom with respect to the bond length. Differences in the location of the CPs provided by different functionals are tiny. The sensitiveness of different indices to the description provided by different functionals is shown in the table. It is seen, for instance, that GGA and global hybrid functionals describe a higher density in the O−H covalent bond than LDA and meta-GGA functionals and, conversely, a lower density in the region of the O$\cdots$H hydrogen bond.

IV. CONCLUSIONS

A parallel implementation in the CRYSTAL program has been presented of a series of algorithms related to the electron density analysis of large molecular and periodic systems. The *a posteriori* analysis of the wave-function can become a computationally demanding task, particularly so if sophisticated techniques are used to extract as much chemical information as possible out of the wave-function of large systems, such as the topological analysis or the high-resolution 3D graphical representation on a dense grid of points. Electron charge and momentum densities, electrostatic potential, X-ray structure factors, Compton profiles and Bader's topological indices can now be computed even for large systems (as large as a crystallized crambin protein which is used to document the efficiency of the speedup of the current implementation) at reduced computational cost for several DFT functionals belonging to four levels of approximation.

Acknowledgments

Carlo Gatti is kindly acknowledged for his continuous support on the merge of the TOPOND package into the CRYSTAL program. Financial support from the Italian Ministry of the University and Research through the DESCARTES project PRIN-2010BNZ3F2 is acknowledged.

* Electronic address: silvia.casassa@unito.it

[1] Gatti, C. and Macchi, P., Eds., *Modern Charge Density Analysis*, Springer, Berlin, 2011.

[2] Hohenberg, P. and Kohn, W., *Phys. Rev.*, 1964, **136**, B864–B871.

[3] Kohn, W. and Sham, L. J., *Phys. Rev.*, 1965, **140**, A1133.

[4] Hirshfeld, F. L., *Israel Journal of Chemistry*, 1977, **16**(2-3), 198–201.

[5] Mulliken, R. S., *J. Chem. Phys.*, 1955, **23**(10), 1833–1840.

[6] Bultinck, P.; Van Alsenoy, C.; Ayers, P. W. and Carbo-Dorca, R., *J. Chem. Phys.*, 2007, **126**(14).

[7] Vanpoucke, D. E. P.; Bultinck, P. and Van Driessche, I., *J. Comput. Chem.*, 2013, **34**(5), 405–417.

[8] Martin Pendas, A.; Costales, A. and Luana, V., *Phys. Rev. B*, 1997, **55**, 4275–4284.

[9] Pendas, A. M.; Francisco, E.; Blanco, M. and Gatti, C., *Chemistry: A European Journal*, 2007, **13**(33), 9362–9371.

[10] de-la Roza, A. O.; Blanco, M.; Pendas, A. M. and Luana, V., *Computer Physics Communications*, 2009, **180**(1), 157 – 166.

[11] Scemama, A.; Caffarel, M. and Savin, A., *J. Comput. Chem.*, 2007, **28**(1), 442–454.

[12] Causá, M. and Savin, A., *Zeitschrift für anorganische und allgemeine Chemie*, 2011, **637**(7-8), 882–884.

[13] Savin, A.; Silvi, B. and Coionna, F., *Canadian Journal of Chemistry*, 1996, **74**(6), 1088–1096.

[14] Johnson, E. R.; Keinan, S.; Mori-Sanchez, P.; Contreras-Garcia, J.; Cohen, A. J. and Yang, W., *J. Am. Chem. Soc.*, 2010, **132**(18), 6498–6506.

[15] Contreras-Garcia, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N. and Yang, W., *J. Chem. Theory Comput.*, 2011, **7**(3), 625–632.

[16] Bader, R. F. W., *Atoms in Molecules - A Quantum Theory*, Oxford University Press, Oxford, UK, 1990.

[17] Dovesi, R.; Orlando, R.; Erba, A.; Zicovich-Wilson, C. M.; Civalleri, B.; Casassa, S.; Maschio, L.; Ferrabone, M.; De La Pierre, M.; D'Arco, Ph.; Noël, Y.; Causá, M.; Rérat, M. and Kirtman, B., *Int. J. Quantum Chem.*, 2014, **114**, 1287–1317.

[18] Orlando, R.; De La Pierre, M.; Zicovich-Wilson, C. M.; Erba, A. and Dovesi, R., *J. Chem. Phys.*, 2014, **141**, 104108.

[19] Orlando, R.; Delle Piane, M.; Bush, I. J.; Ugliengo, P.; Ferrabone, M. and Dovesi, R., *J. Comput. Chem.*, 2012, **33**, 2276–2284.

[20] Perdew, J. P. and Schmidt, K., *AIP Conference Proceedings*, 2001, **577**(1), 1–20.

[21] Gatti, C.; Saunders, V. R. and Roetti, C., *J. Chem. Phys.*, 1994, **101**, 10686.

[22] Topond-2013 : an electron density topological program for systems periodic in n (n=0-3) dimensions, user's manual. Gatti, C. and Casassa, S.; CNR-CSRSRC, Milano, 2013.

[23] Rickwood, P. C.; Mathias, M. and Siebert, J. C., *Contrib. Mineral. Petrol.*, 1968, **19**, 271–301.

[24] Erba, A.; Mahmoud, A.; Orlando, R. and Dovesi, R., *Phys. Chem. Miner.*, 2014, **41**, 151–160.

[25] Teeter, M. M.; Roe, S. and Heo, N. H., *Journal of Molecular Biology*, 1993, **230**(1), 292 – 311.

[26] Yuan, Y.; Mills, M. J. L. and Popelier, P. L. A., *J. Comput. Chem.*, 2014, **35**(5), 343–359.

[27] Pisani, C.; Dovesi, R. and Roetti, C., *Hartree-Fock Ab Initio Treatment of Crystalline solids*, Vol. 48 of *Lecture Notes in Chemistry Series*, Springer Verlag, Berlin, 1988.

[28] Erba, A.; Ferrabone, M.; Orlando, R. and Dovesi, R., *J. Comput. Chem.*, 2013, **34**, 346.

[29] Erba, A. and Pisani, C., *J. Comput. Chem.*, 2012, **33**, 822.

[30] Pattison, P.; Weyrich, W. and Williams, B. G., *Solid State Commun.*, 1977, **21**, 967.

[31] Erba, A.; Pisani, C.; Casassa, S.; Maschio, L.; Schütz, M. and Usvyat, D., *Phys. Rev. B*, 2010, **81**, 165108.

[32] Erba, A.; Itou, M.; Sakurai, Y.; Yamaki, R.; Ito, M.; Casassa, S.; Maschio, L.; Terentjevs, A. and Pisani, C., *Phys. Rev. B*, 2011, **83**, 125208.

[33] Pisani, C.; Erba, A.; Casassa, S.; Itou, M. and Sakurai, Y., *Phys. Rev. B*, 2011, **84**, 245102.

[34] Gatti, C., *Zeitschrift für Kristallographie*, 2005, **220**, 399.

[35] Banerjee, A.; Adams, N. and Simons, J., *J. Phys. Chem.*, 1985, **89**, 52.

[36] Popelier, P. L. A., *Chem. Phys. Lett.*, 1994, **228**, 160.

[37] Espinosa, E.; Alkorta, I.; Elguero, J. and Molins, E., *J. Chem. Phys.*, 2002, **117**, 5529–5542.

[38] Perdew, J. P. and Zunger, A., *Phys. Rev. B*, 1981, **23**, 5048.

[39] Perdew, J. P.; Burke, K. and Ernzerhof, M., *Phys. Rev. Lett.*, 1996, **77**, 3865.

[40] Perdew, J.; Chevary, J.; Vosko, S.; Jackson, K.; Pederson, M.; Singh, D. and Fiolhais, C., *Phys. Rev. B*, 1992, **46**, 6671.

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

[42] Adamo, C. and Barone, V., *J. Chem. Phys.*, 1999, **110**, 6158.

[43] Zhao, Y. and Truhlar, D., *Theor. Chem. Acc.*, 2008, **120**(1-3), 215–241.

[44] E. Espinosa, E. Molins, C. L., *Chem. Phys. Lett.*, 1998, **285**, 170.