

CRYSTAL14: A Program for the Ab initio Investigation of Crystalline Solids

Roberto Dovesi,^{1,*} Roberto Orlando,¹ Alessandro Erba,¹ Claudio M. Zicovich-Wilson,²
Bartolomeo Civalieri,¹ Silvia Casassa,¹ Lorenzo Maschio,¹ Matteo Ferrabone,¹ Marco De la
Pierre,¹ Philippe D'Arco,^{3,4} Yves Noël,^{3,4} Mauro Causá,⁵ Michel Rérat,⁶ and Bernard Kirtman⁷

¹*Dipartimento di Chimica, and Centre of Excellence NIS (Nanostructured Interfaces and Surfaces),
Università di Torino, via Giuria 5, I-10125 Torino (Italy)*

²*Facultad de Ciencias, Universidad Autónoma del Estado de Morelos,
Av. Universidad, 1001, Col. Chamilpa, 62209 Cuernavaca, Morelos (Mexico)*

³*Sorbonne Universités, UPMC Univ Paris 06, UMR 7193,
Institut des Sciences de la Terre Paris (iSTeP), F-75005 Paris, (France)*

⁴*CNRS, UMR 7193, Institut des Sciences de la Terre Paris (iSTeP), F-75005 Paris, (France)*

⁵*Dipartimento di Ingegneria Chimica, dei Materiali e della Produzione Industriale,
Università di Napoli 'Federico II', Napoli (Italy)*

⁶*Equipe de Chimie Physique, IPREM UMR5254,
Université de Pau et des Pays de l'Adour, 64000 Pau (France)*

⁷*Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106 (USA)*

(Dated: February 12, 2014)

The capabilities of the CRYSTAL14 program are presented, and the improvements made with respect to the previous CRYSTAL09 version discussed. CRYSTAL14 is an *ab initio* code that uses a Gaussian-type basis set: both pseudopotential and all-electron strategies are permitted; the latter is not much more expensive than the former up to the first-second transition metal rows of the periodic table. A variety of density functionals is available, including as an extreme case Hartree-Fock; hybrids of various nature (global, range-separated, double) can be used. In particular, a very efficient implementation of global hybrids, such as popular B3LYP and PBE0 prescriptions, allows for such calculations to be performed at relatively low computational cost. The program can treat on the same grounds 0D (molecules), 1D (polymers), 2D (slabs), as well as 3D (crystals) systems. No spurious three-dimensional periodicity is required for low-dimensional systems as happens when plane-waves are used as a basis set. Symmetry is fully exploited at all steps of the calculation; this permits, for example, to investigate nanotubes of increasing radius at a nearly constant cost (better than linear scaling!) or to perform Self-Consistent-Field (SCF) calculations on fullerenes as large as (10,10), with 6000 atoms, 84000 atomic orbitals and 20 SCF cycles, on a single core in one day. Three versions of the code exist, serial, parallel and massive-parallel. In the second one the most relevant matrices are duplicated whereas in the third one the matrices in reciprocal space are distributed for diagonalization. All the relevant vectors are now dynamically allocated and deallocated after use, making CRYSTAL14 much more agile than the previous version, in which they were statically allocated. The program now fits more easily in low-memory machines (as many supercomputers nowadays are). CRYSTAL14 can be used on parallel machines up to a high number of cores (benchmarks up to 10240 cores are documented) with good scalability, the main limitation remaining the diagonalization step. Many tensorial properties can be evaluated in a fully automated way by using a single input keyword: elastic, piezoelectric, photoelastic, dielectric, as well as first and second hyperpolarizabilities, electric field gradients, Born tensors, etc. Many tools permit a complete analysis of the vibrational properties of crystalline compounds. The infrared and Raman intensities are now computed analytically and related spectra can be generated. Isotopic shifts are easily evaluated, frequencies of only a fragment of a large system computed and nuclear contribution to the dielectric tensor determined. New algorithms have been devised for the investigation of solid solutions and disordered systems. The topological analysis of the electron charge density, according to the Quantum Theory of Atoms in Molecules, is now incorporated in the code via the integrated merge of the TOPOND package. Electron correlation can be evaluated at the Möller-Plesset second order level (namely MP2) and a set of double-hybrids are presently available via the integrated merge with the CRYSCOR program.

I. INTRODUCTION

Quantum mechanical *ab initio* simulation is rapidly gaining a more important rôle in many scientific communities due to decreasing computational cost, as well as the availability of computer programs of increasing capability and ease of use. In particular, the number of computer codes devoted to periodic systems has been rapidly grow-

ing. CRYSTAL88 was the first such code to be distributed publicly 25 years ago through the Quantum Chemistry Program Exchange.^{1,2} Since then six other releases have followed in 1992, 1995, 1998, 2003, 2006, 2009 and now there is the current one, CRYSTAL14. This computer program can be used to study the properties of many types of compounds characterized by periodicity in one dimension (quasilinear and helical polymers, nanotubes),

two dimensions (monolayers, slabs), or three dimensions (crystals, solid solutions, substitutionally disordered systems). As a limiting case, molecules can also be studied.

Despite the many improvements and generalizations that have been introduced since its first release, the basic aspects of CRYSTAL have remained the same. Thus, this program computes the electronic structure of periodic systems within the Hartree-Fock (originally) and DFT single particle models using Bloch functions. A special feature of the code is that the crystal orbitals are expanded as linear combinations of atom-centered Gaussian-type functions. Powerful screening techniques are employed to exploit real space locality, which is another distinguishing characteristic of CRYSTAL. Restricted (closed shell) and unrestricted (spin-polarized) calculations can be performed with all-electron or with valence-only basis sets using effective core pseudo-potentials.

Another unique feature is the extensive exploitation of symmetry to achieve computational efficiency. Besides the 230 space groups, 80 two-sided plane groups, 99 rod groups, and 32 crystallographic point groups, there is provision for molecular point group symmetry (e.g. icosahedral) as well as helical symmetry. Automatic tools allow users to obtain lower dimensionality systems from 3D structures by specification of a few geometrical parameters. Slabs (2D periodic), nanorods (1D periodic) and nanocrystals (0D) are easily generated from 3D crystalline structures; nanotubes (1D) and fullerenes (0D)

can be constructed from 2D sheets or multi-layered slabs (Section II A). Full use of symmetry involves all steps of a CRYSTAL calculation, leading to drastically reduced computation time and required memory, as well as improved task farming in parallel calculations (Section II).³⁻⁷ Symmetry is also used to select the independent elements of tensor properties for computation.

Several algorithms of CRYSTAL14 now rely on the numerical accuracy of the geometry optimizer: search for equilibrium structures and transition states, volume- or pressure-constrained minimizations for the determination of the equation of state of bulk crystals, nuclear relaxation of strained lattices for the computation of elastic, piezoelectric, photoelastic tensors, etc. Recent developments of the algorithms of the geometry optimizer are illustrated in Section III.

A wide variety of crystal properties can now be computed automatically. They include third and fourth rank, as well as first and second rank tensors. Amongst the former are the fourth rank elastic tensor along with the related seismic wave velocities;⁸⁻¹¹ the third rank direct and converse piezoelectric tensors,¹² and the fourth rank photoelastic Pockels' tensor (Section IV).¹³ In addition to the dielectric (or polarizability) tensor (static and frequency-dependent), the second- and third-order electric susceptibilities (or third- and fourth-order hyperpolarizabilities) are computed analytically via the Coupled Perturbed Hartree-Fock/Kohn-Sham (CPHF/KS) method (Section V).¹⁴⁻¹⁹

TABLE I: Physical properties that can be computed with CRYSTAL14. For each property, its formula and tensor rank are given along with the general method of computation, which may be either analytical (A) or involve a mixed analytical/numerical (A/N) scheme. Here, $t, u, v, w = x, y, z$ represent Cartesian indices.

Property	Tensor rank	Formula	Definitions	Method
atomic gradient	1	$g_t^a = \frac{\partial E}{\partial r_{a,t}}$	Energy E and atomic position vector \mathbf{r}_a , of atom a ;	A
cell gradient	2	$G_{tu} = \frac{\partial E}{\partial A_{tu}^2}$	Energy E and lattice metric matrix $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$;	A
polarizability	2	$\alpha_{tu} = \frac{\partial^2 E}{\partial \epsilon_t \partial \epsilon_u}$	Energy E and electric field ϵ ;	A
Born charge	2	$Z_{tu}^{a*} = \frac{\partial^2 E}{\partial \epsilon_t \partial r_{a,u}}$	Energy E , electric field ϵ and atomic position vector \mathbf{r}_a ;	A
electric field gradient	2	$q_{tu} = \frac{\partial \epsilon_t}{\partial r_u}$	Electric field ϵ and position vector \mathbf{r} ;	A
Hessian	2	$H_{tu}^{ab} = \frac{\partial^2 E}{\partial r_{a,t} \partial r_{b,u}}$	Energy E and atomic position vector \mathbf{r}_a ;	A/N
direct piezoelectricity	3	$e_{tuv} = \frac{\partial P_t}{\partial \eta_{uv}}$	Polarization \mathbf{P} and rank-2 strain tensor η ;	A/N
converse piezoelectricity	3	$d_{tuv} = \frac{\partial \eta_{uv}}{\partial \epsilon_t}$	Electric field ϵ and rank-2 strain tensor η ;	A/N
first hyperpolarizability	3	$\beta_{tuv} = \frac{\partial^3 E}{\partial \epsilon_t \partial \epsilon_u \partial \epsilon_v}$	Energy E and electric field ϵ ;	A
Raman intensity	3	$I_{tuv}^a = \frac{\partial^3 E}{\partial \epsilon_t \partial \epsilon_u \partial r_{a,v}}$	Energy E , electric field ϵ and atomic position vector \mathbf{r}_a ;	A
elasticity	4	$C_{tuvw} = \frac{\partial^2 E}{\partial \eta_{tu} \partial \eta_{vw}}$	Energy E and rank-2 strain tensor η ;	A/N
photoelasticity	4	$p_{tuvw} = \frac{\partial \Delta \epsilon_{tu}^{-1}}{\partial \eta_{vw}}$	Rank-2 dielectric tensor ϵ and rank-2 strain tensor η ;	A/N
second hyperpolarizability	4	$\gamma_{tuvw} = \frac{\partial^4 E}{\partial \epsilon_t \partial \epsilon_u \partial \epsilon_v \partial \epsilon_w}$	Energy E and electric field ϵ ;	A

Previous versions of the program, starting with CRYSTAL03,^{20,21} have included efficient computation of vibration frequencies (Hessian matrix) and related properties. Now we have added analytical calculation of infrared (IR) and Raman intensities through the CPHF/KS scheme^{22–25} and automated computation of IR and Raman spectra.^{26–28} These vibrational properties are listed in Table I, together with those discussed in the previous paragraph, as the main tensor properties available in CRYSTAL14. In addition, the program now contains improved algorithms for calculating phonon dispersion and anisotropic displacement parameters (ADP)^{29,30} (see Section VI).

New algorithms have recently been developed for the study of solid solutions and, more generally, disordered systems.^{31,32} As far as solid state solutions are concerned, for any substitution fraction x within a given series, the program finds the total number of atomic configurations and determines the symmetry-irreducible configurations among them. Symmetry irreducible configurations can, then, be explored (*i.e.* optimized) automatically. Section VII is devoted to the presentation of these new tools.

We have also developed new tools for the analysis of electron charge and momentum densities. Section VIII deals with i) topological analysis of the electron charge density, performed according to Bader’s prescriptions using Gatti’s TOPOND package,³³ that is now integrated with the CRYSTAL14 program; ii) computation of ADPs and Debye-Waller thermal damping for X-ray structure factors,²⁹ and iii) new algorithms for analyzing the electron momentum density.^{34,35}

For DFT calculations, the local-density and generalized-gradient approximations (LDA and GGA) already available in previous versions of the program have been augmented with exchange-correlation functionals corresponding to the third, fourth and fifth rung of “Jacob’s ladder”. The semilocal and hybrid meta-GGA functionals, range-separated hybrids and double hybrids that have been implemented are discussed in Section IX.

CRYSTAL may be run either sequentially on a single processor or in parallel. Parallel processing, in turn, can be done either through a replicated data procedure (PCRYSTAL), wherein all the most relevant quantities are copied on each node, or according to a MPP (massively parallel) strategy, in which the large matrices are partitioned and distributed amongst the cores. The MP-PCRYSTAL version of the program, first released in 2010, is advantageous for systems with a large unit cell and low symmetry. Since then, we have improved performance so that calculations now scale linearly up to thousands of cores. These recent advances are described in Section X.

The paper is organized as follows: each section corresponds to a particular capability of the CRYSTAL14 program; the newly developed features are illustrated from a general point of view and a few examples are given of their application to systems of interest. Technical details about how to run specific calculations and extract the

corresponding information (input/output structure) can be found in the *CRYSTAL14 User’s Manual*³⁶ and in the many tutorials available at the CRYSTAL website.³⁷

II. USE OF SYMMETRY IN CRYSTAL

Symmetry plays a crucial rôle in the study of crystalline compounds and its use in simulation can be enormously advantageous because: i) it greatly simplifies the set of data to be given as input; ii) it improves performance dramatically; iii) the amount of memory storage is reduced significantly; and iv) it aids in comparing with experiment.

The data needed to define a crystalline structure is greatly reduced if its space group is known because the unit cell can be generated automatically from knowledge of the asymmetric unit only. The same rule applies to slabs, rods, nanotubes and molecules (including fullerenes). Two significant examples are nanotubes and fullerenes, as obtained by geometrical construction from graphene (see Section II A for further details and examples). The same kind of simplification applies to many structure manipulations that modify either translational or point symmetry, or both. For instance, a 2D graphene monolayer can be automatically cut out of a 3D graphite crystal simply by using the keyword SLABCUT (see below).

A. Low Dimensionality Systems

Twenty or thirty years ago, theoretical studies could be classified into three main categories: standard molecular calculations, fully periodic (bulk) calculations, and cluster calculations. The first and last kind of calculations were performed with “molecular” codes, the second kind with “solid state” codes, the two using completely different tools and technologies: plane waves *vs* atomic orbitals, DFT *vs* HF, all-electron *vs* pseudopotential basis set. CRYSTAL, on the contrary, maintains full compatibility along the 3D \rightarrow 2D \rightarrow 1D \rightarrow 0D periodicity series. At the same time, various tools have been implemented for an automated, consistent, easy construction of low-dimensional systems, such as:

Slabs (3D \rightarrow 2D) - This option allows for cutting a 2D periodic slab out of a 3D bulk structure, like a graphene monolayer out of the 3D graphite crystal simply by specifying: i) the crystallographic Miller’s indices of the cutting plane (0 0 1 in this case); ii) the label of the layer corresponding to the bottom cut (any layer in graphite); iii) the number of layers forming the slab (1 in the case of graphene). The origin of the 2D unit cell is set automatically so as to maximize symmetry. In principle, this option allows for any termination of the surface, any stoichiometry and slab polarity because intrinsically “polar” surfaces exist, where the electrical dipole perpendicular to the surface can only be removed by altering the surface structure, for instance, by breaking stoichiometry.

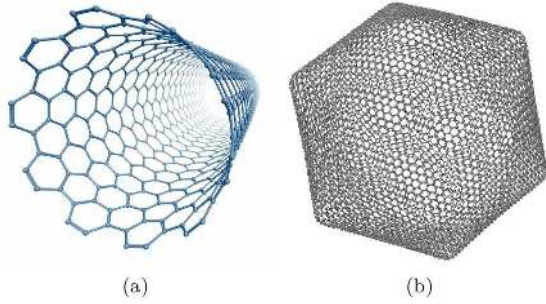


FIG. 1: Structure of (a) a single-walled carbon nanotube and (b) the (10,10) icosahedral carbon fullerene, made up of 6000 atoms.

Nanotubes (2D \rightarrow 1D) - Nanotubes (see Figure 1 a) of any size and order of symmetry can be built automatically by rolling up a 2D slab.³⁸ That is done by specifying a pair of integers to define a roll-up vector in terms of the slab unit vector components. The rolling vector is perpendicular to the nanotube axis and its modulus is the nanotube circumference. In order to build the (120,0) carbon nanotube from a graphene sheet, for instance, the value of those two integers is: 120 and 0.

Fullerenes (2D \rightarrow 0D) - A fullerene cage can be obtained starting from any hexagonal sheet;³⁹ available shapes are: icosahedron, octahedron and tetrahedron (all consisting of equilateral triangular faces). The required information includes a pair of integers representing the components of 2D unit cell vectors defining the side of the triangular face, the fullerene point group and polyhedron type. For example, a (10,10) icosahedral fullerene (see Figure 1 b) is obtained from a graphene sheet by specifying the following data: 10, 10, IH, ICOSA (standing for icosahedron).

Nanorods (3D \rightarrow 1D) - To build “crystals in 1D” starting from the corresponding 3D bulk structure. Nanorods can be used as models for crystal edges (relevant to catalysis), and real nano-objects. The required input steps to build a nanorod are similar to those used to create slabs: information about the crystal plane, termination and thickness of the two crystal planes defining it. At this stage, a nanorod can exhibit sharp and unphysical edges that can be smoothed out by further cutting the rod along additional crystal planes parallel to the periodic direction. The origin of the nanorod cell is shifted automatically to maximize symmetry of the rod group. See Figure 2 (a) for an example.

Nanocrystals (3D \rightarrow 0D) - Non-periodic nanosystems are defined as nanocrystals, provided they preserve a crystalline structure. Very small nanosystems losing the parent crystal structure and stoichiometry are commonly referred to as nanoclusters. A new feature in CRYSTAL14 allows users to create nanocrystals from bulk in a controlled way by generating a supercell with faces parallel (2 by 2) to a set of three crystal planes (see Figure 2 (b) for an example). The same set of information

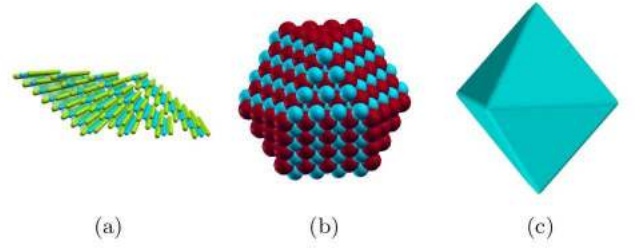


FIG. 2: (a) MgCl_2 nanorod built on the (104) and (001) surfaces;⁴⁰ (b) the $\text{Mg}_{216}\text{O}_{216}$ nanocrystal defined by the (100), (010) and (001) surfaces of bulk MgO, as further edited by the (111) family of planes; (c) Wulff polyhedron of CeO_2 : only (111) and (331) surfaces are represented, with contributions to the total area of 86% and 14%, respectively. Surface energies are reported in Ref. 41.

used to create slabs (and nanorods) needs to be replicated three times in this case. As for nanorods, a nanocrystal can be smoothed by cutting edges along other crystal planes, so breaking the original chemical stoichiometry. For this reason, chemical composition is calculated and printed at each step of the procedure. The origin of the nanocrystal is shifted automatically to maximize point symmetry.

Wulff polyhedron construction (2D \rightarrow 0D) - After a systematic study of a selected set of crystal surfaces, where a set of well converged surface formation energies is obtained, the thermodynamic equilibrium crystal shape can be calculated using the Gibbs-Wulff theory to build a Wulff polyhedron.

B. Symmetry and efficiency

Use of translation invariance is mandatory to transform a cluster calculation for a piece of a bulk system into a periodic bulk calculation. Bloch Functions (BFs) are employed, in that regard, to transform an infinite Hamiltonian matrix into block-diagonal form. Each block corresponds to a \mathbf{k} point in the first Brillouin Zone (BZ) (see the first two panels in Figure 3). All periodic codes are based on the use of BFs.

On the other hand, use of point symmetry is not essential to make a periodic calculation affordable and, indeed, most periodic codes neglect it. However, point symmetry can be very advantageous in periodic calculations (much more than in molecular codes). It is exhibited by most crystalline systems and can be invoked at several steps of an *ab initio* calculation, thereby drastically reducing computation time and resources. The key steps where savings occurs are given in the following for the case where the BFs are constructed from a local basis:

- Diagonalization of the Fock matrix is restricted to the subset of \mathbf{k} points in the irreducible BZ. The eigenvalues in a star of \mathbf{k} points (a set of points

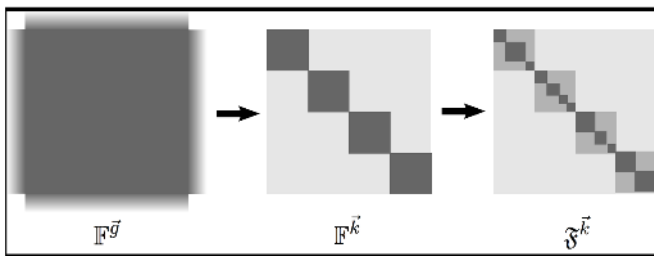


FIG. 3: Block-factorization of the Fock matrix for periodic systems. \mathbb{F}^g : in the basis of AOs, non-packed form (borders are blurry to indicate that such matrix is infinite in principle); \mathbb{F}^k : in the basis of BF's orbitals; \mathbb{F}^k : in the basis of SACOs. Also in the last two cases the matrix is infinite, but block diagonal

that are symmetry related) are the same and the eigenvectors can be generated by symmetry operations. This kind of symmetry is used in CRYSTAL since its first release (CRYSTAL88).

- Time required for the calculation of one- and two-electron integrals is reduced by a factor of up to the number of point symmetry operators in the group. Again, such use of symmetry dates back to CRYSTAL88.³
- Diagonalization of the Fock matrix can be speeded up dramatically in the basis of the Symmetry Adapted Crystalline Orbitals (SACO)/Symmetry Adapted Molecular Orbitals (SAMO), shown in the last panel of Figure 3). SACO/SAMO are generated automatically in CRYSTAL from the selected basis set of Atomic Orbitals (AO) in the unit cell, with no need for additional information about irreducible representations (irrep) or characters. This part of the code was implemented about 15 years ago.^{4,5} The savings factor in computation time is roughly proportional to the third power of the ratio between the number of AOs in the basis set (N_{AO}) and the size of the largest block in the Fock matrix, when represented in the SACO/SAMO basis.
- Construction of the density matrix scales with the third power of the basis set size (N_{AO}^3), as each of the N_{AO}^2 matrix elements is obtained by summing over all occupied crystalline orbitals (very roughly N_{AO}). The advantage is obtained by building the matrix in the SACO/SAMO basis first.

Reduction of computing time thanks to the use of symmetry is not the only issue when handling very large unit cell cases. Memory requirements can also become a bottleneck, if not properly managed at every step of a calculation. Storage of the Fock, overlap and density matrices as full square matrices in the AO-Bloch function basis represent the main bottleneck and need to be avoided. Since both one- and two-electron integrals are

evaluated in the AO basis, a set of back and forth transformations (SACO to AO and AO to SACO) is required. However, it is possible to switch directly from the AO to the SACO/SAMO basis and use the latter throughout. This has been implemented in CRYSTAL14.^{6,7} with a drastic reduction of both running time and memory allocation. The larger the point group size, the bigger the reduction of computational resources (see Section X for details).

In the following we document the savings from full use of symmetry in the case of Carbon Nanotubes (CNT), in particular for the $(n,0)$ family. The number of symmetry operators in CNTs increases with the tube size, i.e. with n . Thus, in principle, there is no limit to the number of non-purely translational elements. Neglecting mirror planes parallel and perpendicular to the tube axis, $(n,0)$ nanotubes possess $2n$ roto-translation symmetry operators. The largest nanotube considered here, with $n = 320$, has 640 such operators and 1280 atoms in its unit cell. This is a huge number compared to a maximum of 48 point operators for standard cubic crystalline systems or 120 for fullerenes, in the molecular context. Thus, large nanotubes are expected to show the maximum symmetry savings factor in terms of computation time and memory allocation. They also represent a severe test for the flexibility of our code and its algorithms, as will be discussed below. A more detailed description of the effect of symmetry can be found in Refs. 6 and 7. Large nanotubes are, clearly, an extreme case. Evidence of the efficiency of the code for systems with lower point symmetry (from 1 to 48 operators) is given in one of the next sections.

Tables II and III show the effect of full use of symmetry on peak memory usage and on running time. The various steps of the calculation can be divided into three units (corresponding to the three blocks in Tables II and III):

- Preliminary calculations for symmetry analysis (including construction of the character table and SACOs), index mapping and screening of integrals (*init*). This unit includes also the generation of a set of orthogonal functions through Cholesky decomposition.
- All steps involved in the SCF cycle, which are iterated as many times as necessary to achieve convergence (15 iterations in the present case). In the “direct SCF” strategy this unit includes: calculation of 1- and 2-electron integrals, as well as integrals approximated by multipolar expansion of the electron charge density (*pole*); transformation of the Fock matrix from the AO to the SACO basis and its diagonalization (*Fock*) + (*diag*); construction of the density matrix in the AO basis (*dens*); and integration of the exchange-correlation density functional (*dft*).
- Calculation of the total energy gradients with respect to the nuclear coordinates (*grad*) at the end of the SCF process.

TABLE II: Peak memory request (in MByte) when one core is used in the SCF+Gradient calculation for a set of $(n,0)$ carbon nanotubes. The B3LYP hybrid functional is used with a basis set of 14 AOs per atom. Only roto-translational operators are included in the point group (mirror planes are neglected). For each nanotube, the number of atoms N_{at} and number of symmetry operators N_{op} is given. The various steps are as follows: initialization of the simulation, including construction of the symmetry group and transformation matrices (*init*); construction of the basis orthogonalization matrix (*Cholesky*); calculation of 1-electron integrals, 2-electron integrals, and integrals approximated by multipole expansion of the charge density (*pole*); transformation of the Fock matrix from the AO to the SACO basis (*Fock*) and its diagonalization (*diag*); construction and back transformation of the density matrix (*dens*); grid integration of the exchange-correlation density functional (*dft*); and calculation of total energy gradients with respect to nuclear coordinates (*grad*).

	(10,0)	(20,0)	(40,0)	(80,0)	(160,0)	(320,0)
<i>init</i>	19	20	45	165	635	2488
<i>Cholesky</i>	17	17	21	32	53	105
<i>2-electron</i>	21	23	30	47	81	138
<i>1-electron</i>	17	18	21	32	53	105
<i>pole</i>	18	18	22	33	55	108
<i>fock + diag</i>	17	18	22	32	54	106
<i>dens*</i>	17	18	22	32	53	105
	(23)	(27)	(40)	(68)	(126)	(250)
<i>dft</i>	19	20	24	35	57	109
<i>grad</i>	74	77	91	260	124	239
N_{at}	40	80	160	320	640	1280
N_{op}	20	40	80	160	320	640

*Formulation of the Density Matrix directly in terms of SACOs will be available in a next release to be distributed in mid 2014. Data in parentheses refer to the release distributed since December 2013.

Let us begin with the memory allocation requirements reported in Table II. Clearly, all steps except *init* require an extremely small amount of memory. As an example, the 239 MBytes for the *grad* step in the case of the largest nanotube is about 1/10 of the memory commonly available in a single core standard machine. The reason for such low memory requirements is that the number of atoms in the asymmetric unit is just 2 regardless of n . In CRYSTAL14 only the irreducible part of the Fock, overlap and density matrices is stored in memory (we recall that in addition, screening reduces the total number of matrix elements dramatically). As a consequence, about 6000 elements are stored for each matrix, irrespectively of the size of the nanotube.

In the *init* step, the generation of the coefficients for SACOs has not yet been optimized from the viewpoint of memory allocation, as this is not an issue in all normal

TABLE III: CPU time (in seconds) for the various steps in the SCF+Grad calculation for $(n,0)$ carbon nanotubes. See Table II for description of individual steps. Timing for a single SCF cycle (TOT_{cyc}) and for the entire SCF procedure (TOT_{SCF} , 15 cycles in this case) are also reported.

	(10,0)	(20,0)	(40,0)	(80,0)	(160,0)	(320,0)
<i>init</i>	1.70	16.00	10.38	71.37	602.51	6461.87
<i>Cholesky</i>	0.59	2.17	8.55	34.63	138.89	561.94
<i>pole</i>	0.24	0.47	0.94	1.90	3.80	7.60
<i>2-electron</i>	26.29	21.39	20.72	21.28	22.45	34.84
<i>1-electron</i>	0.20	0.34	0.68	1.31	2.58	5.24
<i>Fock</i>	0.44	1.59	7.17	30.85	128.36	525.69
<i>diag</i>	0.10	0.39	0.81	1.58	3.07	6.67
<i>dens*</i>	0.05	0.09	0.18	0.38	0.90	2.12
	(0.92)	(3.28)	(13.38)	(54.83)	(221.80)	(883.44)
<i>dft</i>	11.52	11.61	13.14	15.17	20.65	34.80
TOT_{cyc}	38.88	36.00	43.86	73.06	182.68	619.11
TOT_{SCF}	585.49	558.17	676.83	1201.90	3481.60	16310.46
<i>grad</i>	240.52	203.72	208.83	249.45	388.23	904.11

cases (any space group or icosahedral symmetry). However, because memory requirements increase quadratically with the number of operators, SACO generation becomes dominant at about 150-200 operators. In the case with the largest number of symmetry operators (640) considered here more than 2 GBytes are required. That is why only roto-translation operators were taken into account for Table II. The addition of mirror planes would quadruple the number of symmetry operators, and thereby increase memory requirements, with little gain in efficiency.

Let us now turn to the running time. Table III shows that the full SCF procedure plus gradient calculation for the largest CNT (1280 atoms, 17920 atomic orbitals per unit cell) is accomplished within less than 5 hours on a single processor for the B3LYP hybrid functional. No calculation for such a large CNT has been reported in the literature, even with smaller basis sets or simpler functionals such as LDA or GGA. The cost of the various steps is reported in Table III and represented pictorially in Figure 4; the (expectedly) most demanding steps along the series are 2-electron integrals and calculation of the gradients (*grad*). Nevertheless, for the same reasons invoked for low memory requirement, computing time is remarkably little and nearly independent of the nanotube size: quite interestingly, about 40% of computing time for the largest tube is spent in the *init*, *Cholesky* and *Fock* steps, where higher efficiency could, in principle, be recovered.

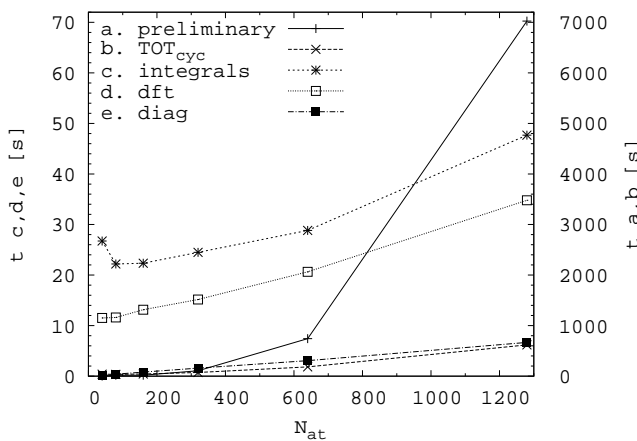


FIG. 4: CPU time (in seconds) for the various steps of the calculation of carbon nanotubes. See Table II for definitions. TOT_{cyc} is the time required for a SCF cycle at B3LYP level.

C. Symmetry in anisotropic properties

A number of physical properties can be given a tensorial representation. For example, linear elasticity is described by a fourth order tensor. Such a tensor consists of 21 independent components for a triclinic crystal whereas they reduce to 3 in the case of a cubic crystal. Thus, symmetry is key to study tensorial properties: which components have to be computed? Which elements are null by symmetry? Which are symmetry-related? Moreover, experimental data are generally reported with some standard orientation of the cell parameters with respect to the Cartesian frame. For comparison between calculated and experimental data to be consistent, the orientation issue needs to be clearly stated. CRYSTAL performs such a symmetry analysis of the tensor using the **TENSOR** keyword. That is particularly useful in low-symmetric cases and in the case of three-fold rotation axes. Such analysis is automatically performed prior to the calculation of all tensorial properties available. A single keyword (**ELASTCON** for elastic constants, **PIEZOCON** for piezoelectric constants, **PHOTOELA** for photoelastic constants, see Section IV for details) is sufficient for generating the full tensor of interest.

III. GEOMETRY OPTIMIZATION OF PERIODIC SYSTEMS

Geometry optimization is mostly employed in quantum chemical calculations to obtain a nuclear configuration that is either

- a stable structure of a given chemical species, or
- can be used to estimate the transition configuration along a reaction path leading to determination of the rate for the corresponding process.

The former use is based on the Born-Oppenheimer approximation and thermostatics, and is realized by computing the corresponding minimum of the Potential Energy Hypersurface (PEH). As concerns the latter, an optimization towards an appropriate saddle point of the PEH is done and, then, thermostatical considerations yield activation energies and reaction rates. These are the main reasons why geometry optimization is one of the most used options of any quantum chemical code. In what follows, the optimization features recently implemented in CRYSTAL14, mostly dealing with minimization, will be briefly discussed.

The efficiency of minimization methods depends primarily on the extent to which the energy function or PEH can be represented by a pure quadratic power series expansion in terms of the geometric coordinates used. Most gradient techniques warrant that, for an ideal pure quadratic PEH, the minimum will be attained in less than M steps, where M is the number of independent coordinates.⁴² Normally, such ideal behavior is not found and various techniques are employed to ensure similar convergence behavior.

In the CRYSTAL code two different strategies are utilized to deal with a non-quadratic periodic PEH: the *trust radius* technique and the choice of suitable coordinates. The former has already been well documented,⁴³ and since no special features are used in the code implementation, it will not be considered further. On the other hand, the coordinate system choices available in CRYSTAL for periodic calculations are worthy of discussion.

Although the critical points of a PEH are invariant under coordinate system transformations, the quadratic character of the energy function may be drastically improved with a proper geometrical parameter choice. CRYSTAL allows for the use of two coordinate systems that are particularly suitable for periodic calculations, namely symmetrized crystallographic (SC) and redundant valence (RV) coordinates. The former is the basic set for any crystalline system and consists of symmetry adapted fractional directions (SADIR) and elastic deformations (SAED). In CRYSTAL14 both sets of geometrical parameters are normalized by default. This ensures equivalent weights in the optimization procedure, which is particularly relevant for the construction and use of the approximate Hessian matrix employed in *pseudo*-Newton techniques.⁴²

With regard to SADIRs, the translational degrees of freedom are first excluded. Then, the remaining directional coordinates are symmetry adapted and, subsequently orthonormalized using a Schmidt procedure. For SAEDs the set is first symmetry adapted and Schmidt orthogonalized. Then, each deformation d is normalized so as to fulfill the condition

$$\sum_{A=1}^N \sum_{t=1}^3 \left[\sum_{u=1}^3 \epsilon_{tu}^{(d)} X_u^A \right]^2 = 1, \quad (1)$$

where N is the number of atoms, X_u^A are the Cartesian coordinates of atom A and $\epsilon_{tu}^{(d)}$ is the d -th SAED.

While the SC coordinates are easily understood in terms of common crystallographic concepts, their main drawback is that the resulting energy function is frequently far from being pure quadratic. There is also, typically, a strong coupling between the atomic directional and elastic deformation degrees of freedom. This is particularly noticeable in covalent systems that exhibit strongly directed bonds. As previously proposed,^{44,45} for this kind of system the use of RV coordinates often improves the optimization process.

The generation of the RV sets and their implementation for geometry optimization of molecules^{46,47} and crystals^{48,49} has been well-described in the previous literature. Accordingly, we omit the details here and highlight just the key differences in CRYSTAL from other implementations.

The atomic connectivity required to define the RV coordinates is set up according to Refs. 47 and 46. An additional feature in the present implementation is that symmetry equivalences within the RV set are recognized and only one representative of each class is kept in memory together with the corresponding multiplicity per unit cell, μ_i . A small displacement in the reference coordinate system, $\delta\mathbf{x}$, can be transformed into the displacement $\delta\mathbf{q}$ in the RV basis set using

$$\delta\mathbf{q} = \mathbf{B}\delta\mathbf{x}, \quad (2)$$

where \mathbf{B} is the Wilson B-matrix whose elements are given by

$$B_{ij} = \mu_i \partial q_i / \partial x_j.$$

In CRYSTAL the reference coordinate system, \mathbf{x} consists of SADIRs and SAEDs and the B-matrix is computed by numerical differentiation based on the central point approximation. The force in RV coordinates is determined from the force in the reference system according to

$$\mathbf{f}_q = \mathbf{B}^- \mathbf{f}_x, \quad (3)$$

where $f_{q_i} = -\partial E / \partial q_i$ and $\mathbf{B}^- = \mathbf{G}^- \mathbf{B}^T$, the superscript “-” indicates the generalized inverse and $\mathbf{G} = \mathbf{B}^T \mathbf{B}$.

While optimization is carried out in the redundant space, in some steps of the procedure vectors in spaces tangent to the PEH must be computed. This is done by employing the projector $\mathbf{P} = \mathbf{G}^- \mathbf{G}$.⁴⁷ If constraints are required to keep some RV parameters fixed, then a new projector that eliminates redundancies as well, \mathbf{P}' , is constructed as⁴⁶

$$\mathbf{P}' = \mathbf{P} - \mathbf{P}\mathbf{C}(\mathbf{C}\mathbf{P})^- \mathbf{C}\mathbf{P}, \quad (4)$$

where \mathbf{C} is the projector onto the constraints subspace, which is given in RV coordinates by

$$\mathbf{C} \equiv C_{ij} = \begin{cases} 1 & \text{if } i = j \text{ and } i \text{ is constrained,} \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Both gradient and Hessian have to be projected. With regard to the latter, the projected matrix $\tilde{\mathbf{H}} = \mathbf{P}'\mathbf{H}\mathbf{P}'$ is diagonalized and its generalized inverse computed according to

$$[\tilde{H}_{ij}]^- = \sum_k T_{ik} [h_k]^{-1} T_{jk}, \quad (6)$$

where T_{ik} is an element of the eigenvector matrix and k runs over all eigenvalues of $\tilde{\mathbf{H}}$ except those that correspond to the redundancies (i.e. for n redundancies the n eigenvalues of smallest magnitude). This procedure also corrects for small errors that derive from numerical calculation of the B-matrix.

Once the optimized displacements in the RV coordinate system have been found, they must be back-transformed to the reference system to obtain the resulting geometry. In CRYSTAL14 this is performed two different ways:

1. in an iterative manner similar to that proposed for molecular codes,⁴⁶ and
2. by a strategy that minimizes a function which measures the proximity of an RV set of coordinates to a real crystal geometry.

Option 1 has proved to be efficient for most molecular structures. However, it is not as safe when points in the RV are very far from the subspace that matches the true geometrical parameter space. This usually occurs when the connectivity that defines the RV coordinates exhibits a large number of nested loops, which is the case in most periodic systems, and the required change in geometry is large. In such cases, the iterative procedure of option 1 diverges and a very approximate solution of the back-transformation relation must be chosen in order to continue the optimization.⁴⁷ When such an approximate back-transformation is performed more than once during the optimization, the inaccuracy of the displacements usually makes the whole procedure inefficient and most of the advantages due to adopting RV coordinates is lost.

Option 2 is now implemented as an alternative in CRYSTAL14. It allows a more accurate back-transformation, thereby providing a substantial reduction in the number of optimization steps. The overall concept is as follows. Let us consider a set of RV coordinates $\mathcal{Q} = \{q_i\}$ and a reference set, $\mathcal{X} = \{x_i\}$, whose redundancy is, in general, less than the former. There is a correspondence such that every vector expressed in terms of \mathcal{X} has an image in terms of \mathcal{Q} . The converse is not, in general, true. The aim of the technique discussed here is to unambiguously assign a coordinate position in \mathcal{X} given a point in \mathcal{Q} . Thus, given a displacement in terms of RV coordinates one would like to find a displacement in the reference set that is closest to the former. The idea of “closest” can be quantified by an “error” function:

$$\Phi(\mathbf{x}) = \sum_j a_j (q_j - \bar{q}_j)^2, \quad (7)$$

where a_j are weight factors, \bar{q}_j is the target in RV coordinates and $q_j = q_j(x_1, x_2, \dots, x_N)$. The optimum point in \mathcal{X} space is the one that minimizes Φ .

The gradient of Φ may be written as

$$\frac{\partial \Phi}{\partial x_k} = 2 \sum_j a_j (q_j - \bar{q}_j) \frac{\partial q_j}{\partial x_k}, \quad (8)$$

with $\partial q_j / \partial x_k = B_{jk}$. If we define the vector $d_j = 2a_j(q_j - \bar{q}_j)$, then eq. (8) becomes:

$$\mathbf{g} = \mathbf{B} \cdot \mathbf{d}, \quad (9)$$

where $g_k = \partial \Phi / \partial x_k$. The optimum point in \mathcal{X} , corresponding to zero gradient, is obtained by the conjugate gradient method which, in the present implementation, utilizes the Polak-Ribiere search direction.^{42,50}

IV. ELASTIC, PIEZOELECTRIC AND PHOTOELASTIC PROPERTIES OF CRYSTALS

Several strain-related tensor properties of crystalline materials can be computed with the CRYSTAL14 program simply by inserting a single keyword in the input deck. These properties include elasticity, piezoelectricity and photoelasticity. A general and robust algorithm has been devised which automatically handles any space group and fully exploits point symmetry. Thus, along with other general improvements of the code described in Section X, it is now possible to carry out a complete *ab initio* calculation of elastic properties for large unit cell systems, such as garnets with 80 atoms per cell (see Section IV A below for details).^{9,10}

The elements of the fourth-rank elastic tensor \mathbb{C} are second energy density derivatives with respect to pairs of cell deformations (see Table I). A fully automated algorithm, using analytical energy gradients,⁸ for the calculation of \mathbb{C} , which was already implemented in CRYSTAL09 for 3D systems, has now been extended in a number of ways including the:^{9–11} i) generalization to 1D and 2D systems, ii) calculation of directional seismic wave-velocities, iii) calculation of polycrystalline isotropic aggregate elastic properties, iv) calculation of elastic properties under pressure. Some of these points are discussed in Section IV A.

The direct piezoelectric tensor \mathbf{e} , whose elements are first derivatives of the polarization with respect to crystal strain, is third rank. In our fully automated implementation,¹² the polarization is computed via the Berry phase approach.⁵¹ The converse piezoelectric tensor \mathbf{d} , determined as the strain induced by an external electric field, is evaluated from \mathbf{e} and \mathbb{C} using $\mathbf{d} = \mathbf{e}\mathbb{C}^{-1}$. These calculations can be done for 1D, 2D, 3D systems. A more detailed discussion of piezoelectricity will be given in Section IV B along with a brief review of recent applications.

The elements of Pockels' photoelastic, or elasto-optic, fourth-rank tensor \mathbb{P} , are defined as first derivatives

of the inverse of the dielectric tensor with respect to crystal strain. An automated scheme has been implemented,¹³ which adopts the Coupled-Perturbed-Hartree-Fock/Kohn-Sham (CPHF/KS) method (see Section V) for computing the dielectric tensor of equilibrium and strained configurations. The explicit dependence of elasto-optic constants on the electric field frequency can be evaluated as well, thus allowing a direct comparison with experimental data. Some examples will be given in Section IV C.

A. Elasticity

1. Nanotubes and Monolayers

As noted above, several improvements have recently been made to the algorithms for computing elastic properties of crystals. One of them is the generalization to 1D and 2D systems.¹¹ In this work we have investigated elastic and vibrational properties for several families of single-walled nanotubes and discussed how they are expected to converge to those of the corresponding monolayers. Zigzag ($n,0$) boron nitride,¹¹ zigzag ($n,0$) beryllium oxide,⁵² armchair (n,n) zinc oxide⁵³ and zigzag ($n,0$) magnesium oxide⁵⁴ nanotubes have been studied, with tube radii up to 24 Å, 27.1 Å, 45.3 Å and 43.3 Å, respectively.

A technologically interesting and widely discussed feature of nanotubes is their response to uniaxial strain along the nanotube axis.^{55,56} In this regard reference is made to the Young's (elastic) modulus, which essentially coincides with the C_{11} elastic constant. A comparison with the C_{11} elastic constant of the corresponding monolayer, as the tube radius increases, is not straightforward. As discussed in Ref. 52, when a nanotube is stretched (compressed) in the axial direction, its radius decreases (increases) in order to minimize the total energy. For the monolayer there is a corresponding deformation orthogonal to the applied strain, *i.e.* Poisson's effect, that must be taken into account. The Poisson-corrected monolayer value, for comparison, turns out to be $C_{11} \times (1 - \sigma^2)$, where σ is Poisson's ratio given by $\sigma = -C_{12}/C_{11}$ in terms of monolayer elastic constants. The damping factor $(1 - \sigma^2)$ is almost negligible for graphene and *h*-BN (0.970 and 0.955) whereas it is 0.867 for *h*-BeO and 0.577 for *h*-ZnO. Clearly, this effect increases with the ionicity of the chemical bonds in the system.

The IR-active vibrations of the above nanotubes can be subdivided into two distinct groups, with frequencies ν that tend either to an optical frequency of the monolayer or to zero with increasing tube size n . Three modes decrease linearly to zero as $1/n$. These are collective modes without a direct correspondence in the vibrational spectrum of the monolayer. One of them, for instance, is the ring breathing. We have demonstrated that, in the large radius limit, these modes can be related to elastic deformations of the monolayer.¹¹ By imposing equality

between the elastic strain energy of the monolayer and the corresponding vibrational energy of the nanotube, the slope of frequency versus $1/n$ for the latter can be related to the elastic constants of the monolayer. For the ring breathing mode, this gives:

$$\nu = \sqrt{\frac{C_{22}}{(M_A + M_B) n |\mathbf{a}|}}, \quad (10)$$

where M_A and M_B are the atomic masses of the two atoms of the system (boron and nitrogen, zinc and oxygen, etc.) and \mathbf{a} is the lattice parameter of the monolayer. By fitting the BN nanotube vibration frequency versus $1/n$ we obtain 2637 cm^{-1} for the lhs of Eq. (10), whereas the elastic constants of the monolayer give a value of 2662 cm^{-1} . For BeO the same comparison yields 1913 cm^{-1} versus 1915 cm^{-1} , while for ZnO the two values are 672 cm^{-1} and 668 cm^{-1} . Bearing in mind that the properties involved in the comparison (vibration frequencies of the nanotube and elastic constants of the monolayer) are computed quite differently, the agreement is remarkable and confirms the high accuracy of all the algorithms involved.

2. Seismic Wave Velocities and Elastic Anisotropy

The acoustic wave velocities of a crystal are related to the elastic constants by Christoffel's equation.^{57,58} In CRYSTAL14, an automated procedure has been implemented for computing these wave velocities along any crystallographic direction. The three acoustic wave velocities, also referred to as seismic velocities, can be labeled as quasi-longitudinal v_p , slow quasi-transverse v_{s1} and fast quasi-transverse v_{s2} , depending on their polarization with respect to the propagation direction.⁵⁹

Silicate garnets are among the most important rock-forming minerals; they represent the main constituents of the Earth's lower crust, upper mantle and transition zone. These garnets are characterized by a cubic lattice with space group $Ia\bar{3}d$ and formula $X_3Y_2(\text{SiO}_4)_3$, 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+} . The primitive cell contains four formula units (80 atoms) and the structure consists of alternating SiO_4 tetrahedra and YO_6 octahedra sharing corners to form a three-dimensional network. In a recent work the B3LYP elastic properties, including seismic wave velocities, were obtained for the six most abundant end-members of this family (pyrope, almandine, spessartine, grossular, uvarovite and andradite).^{9,10}

In Figure 5, we compare our *ab initio* directional seismic wave velocities for andradite $\text{Ca}_3\text{Fe}_2(\text{SiO}_4)_3$, in particular, with experimental values.⁶⁰ Seismic wave velocities are reported along an angle θ such that $\theta = 0^\circ$ corresponds to the crystallographic direction [110], $\theta = 45^\circ$ to the [111] direction, $\theta = 90^\circ$ to the [001] direction, etc. The agreement is quite impressive: both the angu-

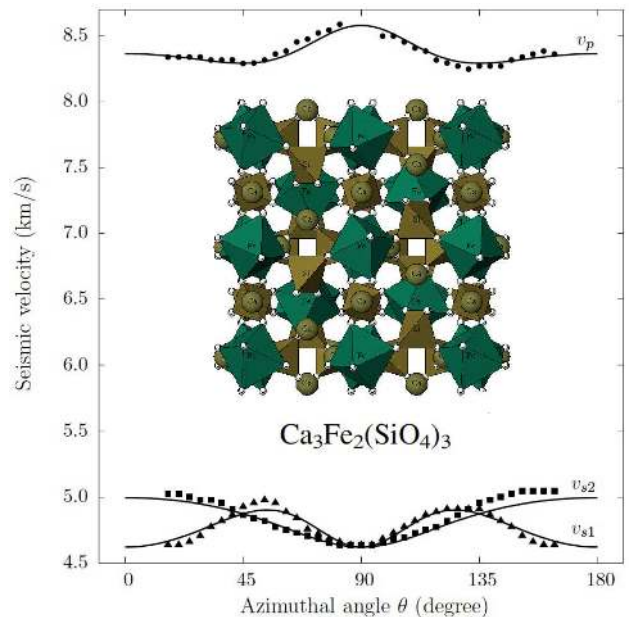


FIG. 5: Directional seismic wave velocities of andradite single-crystal, $\text{Ca}_3\text{Fe}_2(\text{SiO}_4)_3$, as computed at the B3LYP level (continuous lines) and as measured by Brillouin scattering at ambient pressure by Jiang *et al.*⁶⁰ (black symbols). Seismic wave velocities (quasi-longitudinal v_p , slow and fast quasi-transverse v_{s1} and v_{s2}) are reported along an angle θ defined in the text. Computed values have been downshifted by 0.1 km/s.

lar dependence and the oscillation amplitudes are very accurately reproduced.

A further elastic property of great interest is the so-called elastic wave anisotropy, as measured by the dimensionless parameter A that vanishes for an isotropic material. For cubic crystals, A is given by the following simple expression in terms of elastic constants: $A = (2C_{44} + C_{12})/(C_{11}) - 1$. The elastic anisotropy of garnets is generally rather small, 0.6 %, when compared to that of other rock-forming minerals such as olivine, 25 %, spinel, 12 %, muscovite, 58 %, orthopyroxene, 16 %, etc.⁶¹ From the calculations of A , we can sort the six silicate garnet end-members according to their increasing elastic anisotropy: spessartine < pyrope < grossular < almandine < andradite << uvarovite. Spessartine and pyrope show very low anisotropies ($A = -0.025$ and $A = -0.031$, respectively) whereas uvarovite is, by far, the most anisotropic with $A = -0.159$.

B. Piezoelectricity

Direct and converse piezoelectric tensors of 1D, 2D and 3D periodic systems can now be computed automatically with CRYSTAL14.

The piezoelectric response of nanotubes is of considerable interest. We have recently found that BeO nanotubes exhibit a longitudinal piezoelectric response that

is about 25% larger than BN nanotubes of comparable radii.⁵² This is noteworthy since the response of the latter is quite large amongst low-dimensional systems.⁶² In comparing the piezoelectric constant e_{11} for nanotubes of increasing radii with that of the corresponding monolayer, Poisson's effect has, again, to be taken into account. As before (see Sec. IV A) the connection between vibration frequencies of the nanotubes and elasticity of the monolayer implies a relationship, now between the contribution of the nanotube collective modes to the polarizability on one hand and the monolayer piezoelectricity on the other. We have investigated this relationship for BN, BeO and ZnO nanotubes.^{11,52,53}

As regards 3D systems, standard piezoelectric ceramics, such as lead zirconate titanate (PZT) based materials, are widely used in many technological applications.^{63,64} At cryogenic temperatures, however, their piezoelectric response is significantly reduced. Hence, they cannot be used as actuators for adaptive optics in space telescopes and low temperature capacitors. In 1997 Grupp and Goldman discovered a giant piezoelectric effect in strontium titanate SrTiO_3 down to 1.6 K, where the only non-zero converse piezoelectric coefficient $d_{31} = 16 \times 10^{-10}$ m/V was reported. This is comparable to the value for PZT at room temperature.⁶⁵ In two recent studies, we fully characterized direct and converse third-rank piezoelectric tensors of the low temperature phases of SrTiO_3 and BaTiO_3 .^{12,66}

Due to its peculiar piezoelectric properties, $\alpha\text{-SiO}_2$ is another material that is widely utilized in the electronics industry. Unfortunately, its suitability is reduced for applications requiring high thermal stability and high electromechanical coupling. These limitations are mainly due to the $\alpha\text{-SiO}_2$ to $\beta\text{-SiO}_2$ phase transition, in which case the piezoelectric constant d_{11} vanishes and d_{14} remains the only non-zero component.⁶⁷ Among quartz homeotypes, GaAsO_4 and $\alpha\text{-GeO}_2$ exhibit the highest electromechanical coupling coefficients resulting in an electrical to mechanical energy conversion efficiency of about 22 %.⁶⁸ Furthermore, they show a very high degree of thermal stability since they do not undergo an α - β phase transition.⁶⁹ Using the new tools in CRYSTAL14 we have recently studied the solid solution $\text{Si}_{1-x}\text{Ge}_x\text{O}_2$ of α -quartz, where silicon atoms are progressively substituted with germanium atoms, as a function of x . A linear increase in the piezoelectric response with the substitutional fraction had been suggested experimentally for very small x .⁷⁰ Our exploration of the entire range of substitution revealed a quasi-linear increase throughout, a result that provides the basis for an effective tuning of the piezoelectric response.⁷¹

C. Photoelasticity

The variation of the refractive index (dielectric constant) with respect to internal or applied strain gives rise to the so-called photoelastic, or elasto-optic, tensor. We

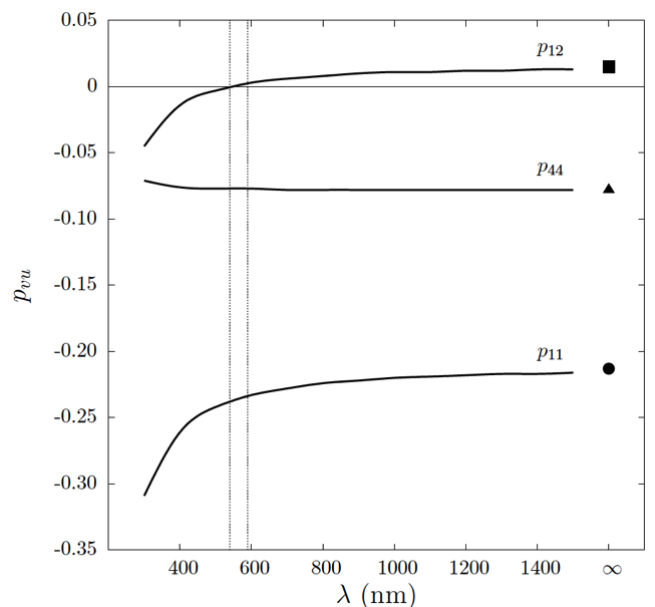


FIG. 6: The three independent elasto-optic constants of MgO crystal, computed at the PBE level, as a function of the electric field wavelength λ . Infinite wavelength results are also shown, while dashed vertical lines indicate the experimental wavelength range.

have recently implemented a general algorithm for the *ab initio* calculation of this fourth-rank tensor in the CRYSTAL14 version of our program.¹³ Using this code we computed the elasto-optic constants for a set of 8 crystalline systems of different symmetry: simple cubic sodium chloride NaCl, lithium fluoride LiF, magnesium oxide MgO and potassium chloride KCl; cubic silicon and diamond; trigonal SiO_2 α -quartz and tetragonal TiO_2 rutile. Good overall agreement with experiment was obtained, particularly so for the PBE generalized-gradient functional.

Experimental results for the elasto-optic constants can vary considerably from one measurement to another. A case in point is MgO, for which there are three independent constants, namely p_{11} , p_{12} and p_{44} . The individual values of p_{11} and p_{12} deduced by several workers are very different from one another: p_{11} ranges from -0.21 to -0.31 and, for p_{12} , even the sign is uncertain, with values ranging from -0.08 to 0.04.⁷²⁻⁷⁴

Although the experiments are supposed to measure the variation of the *static* dielectric tensor (*i.e.* at infinite electric field wavelength, $\lambda = \infty$), they are actually performed at finite wavelengths that may give results far from the static limit. The experimental values reported above were obtained in the wavelength range from 540 nm to 589.3 nm. In Figure 6, we show the three independent elasto-optic constants of MgO, computed at the PBE level, as a function of λ (see Section V B for more details about the λ -dependent CPHF/KS computational scheme). While p_{44} is almost wavelength-independent, p_{11} and p_{12} exhibit a clear dependence upon λ , slowly converging to the static limit above 1000 nm. In partic-

ular, the value of p_{12} is found to pass from negative to positive around 550 nm. Dashed vertical lines in the figure identify the experimental range of electric field wavelengths; both p_{11} and p_{12} are still changing in that range. This aspect is particularly crucial for p_{12} which changes sign in that range. This explains the uncertainty in its experimental value.

As a further application, we have recently characterized the photoelastic behavior of the low-temperature rhombohedral phase of BaTiO₃, again by explicitly treating the dependence on λ .⁶⁶

V. STATIC NON-LINEAR POLARIZABILITIES AND FREQUENCY-DEPENDENT ELECTRONIC DIELECTRIC CONSTANT

A new feature in CRYSTAL14 is the calculation of the electronic and vibrational contributions to the first and second static hyperpolarizability tensor for molecules, polymers, slabs and crystals. The dependence of the electronic linear polarizability (or dielectric constant) on the frequency of the applied electric field has been added as well. The electronic properties are computed analytically, while the vibrational properties require a finite field geometry optimization.

A. Coupled Perturbed HF/KS Calculation of Static Electronic (Hyper)polarizabilities

The total electronic energy E of a crystal in a uniform static electric field ε can be expressed as

$$E(\varepsilon) = E(0) - \sum_t \mu_t \varepsilon_t - \frac{1}{2!} \sum_{tu} \alpha_{tu} \varepsilon_t \varepsilon_u + \dots - \frac{1}{3!} \sum_{tuv} \beta_{tuv} \varepsilon_t \varepsilon_u \varepsilon_v - \frac{1}{4!} \sum_{tuvw} \gamma_{tuvw} \varepsilon_t \varepsilon_u \varepsilon_v \varepsilon_w + \dots$$

with $E(0)$ being the field-free electronic energy and $\mu, \alpha, \beta, \gamma \dots$ the electronic energy derivatives of order

1, 2, 3, 4 *dots* with respect to the Cartesian components of the electric field (indicated by the subscripts t, u, v, w). As for the corresponding physical properties, μ represents the dipole moment, α the polarizability, β the first hyperpolarizability and γ the second hyperpolarizability. For a closed-shell system, the second energy derivatives α_{tu} are calculated using the expression:

$$\alpha_{tu} = -\frac{2}{n_k} \sum_{\mathbf{k}}^{\text{BZ}} \Re \left\{ \text{Tr} \left(C^{\mathbf{k}\dagger} \Omega^{\mathbf{k},t} C^{\mathbf{k}} U^{\mathbf{k},u} n \right) \right\} \quad (11)$$

where n_k is the number of \mathbf{k} points in the first Brillouin Zone (BZ) and n is the diagonal occupation matrix whose elements are 2 for occupied orbitals and 0 otherwise. Here $U^{\mathbf{k},u}$ is an anti-Hermitian block off-diagonal matrix that relates the unperturbed coefficient matrix $C^{\mathbf{k}}$, to the corresponding first-order perturbed matrix, $\mathbb{C}^{\mathbf{k},u} \equiv C^{\mathbf{k}} U^{\mathbf{k},u}$, which gives the linear (first-order) response to the electric field perturbation represented by the matrix $\Omega^{\mathbf{k},t}$. The off-diagonal blocks of $U^{\mathbf{k},u}$ depend not only on $\Omega^{\mathbf{k},u}$ but also on the first-order perturbed density matrix (through the two-electron terms in the Hamiltonian) which, in turn, depends upon $U^{\mathbf{k},u}$. Hence, a self-consistent solution of the Coupled-Perturbed Hartree-Fock equations is required.

The CPHF calculation of static linear polarizabilities was already available in a previous version of CRYSTAL,^{14,15} for both closed and open shell systems, along with a corresponding CPKS treatment using pure and hybrid DFT functionals.¹⁶ Now the CPHF/KS method has been extended to the first and second hyperpolarizability tensors. As a consequence of the 2n+1 rule the first-order coefficient matrix is sufficient to determine the first hyperpolarizability tensor. However, calculation of the second hyperpolarizability tensor γ requires the second-order coefficients and they must be obtained from a self-consistent solution of the second-order coupled perturbed equations. Both the β and γ tensors can be expressed in terms of $U^{\mathbf{k},uv}$:

$$\beta_{tuv} = -\frac{1}{n_k} \sum_{\mathbf{k}} \Re \left\{ \mathcal{P}_{t,u,v} \text{Tr} \left(\left[C^{\mathbf{k}\dagger} \mathbb{F}^{\mathbf{k},t} C^{\mathbf{k}} U^{\mathbf{k},uv} - C^{\mathbf{k}\dagger} \mathbb{F}^{\mathbf{k},uv} C^{\mathbf{k}} U^{\mathbf{k},t} \right] n \right) \right\} \quad (12)$$

$$\begin{aligned} \gamma_{tuvw} = & -\frac{1}{n_k} \sum_{\mathbf{k}} \Re \left\{ \mathcal{P}_{t,u,v,w} \text{Tr} \left(U^{\mathbf{k},t\dagger} \left[C^{\mathbf{k}\dagger} \left(\mathbb{F}^{\mathbf{k},u} C^{\mathbf{k}} U^{\mathbf{k},vw} + \frac{1}{2} \mathbb{F}^{\mathbf{k},vw} C^{\mathbf{k}} U^{\mathbf{k},u} \right) - \right. \right. \right. \\ & \left. \left. \left(U^{\mathbf{k},vw} C^{\mathbf{k}\dagger} \mathbb{F}^{\mathbf{k},u} C^{\mathbf{k}} + \frac{1}{2} U^{\mathbf{k},u} \mathbb{F}^{\mathbf{k},vw} \right) \right] n + i U^{\mathbf{k},tw\dagger} \frac{\partial U^{\mathbf{k},u}}{\partial k_v} n \right) \right\} \end{aligned} \quad (13)$$

The operator \mathcal{P} in Eqs. (12) and (13) carries out the

sum over all permutations of the Cartesian directions and

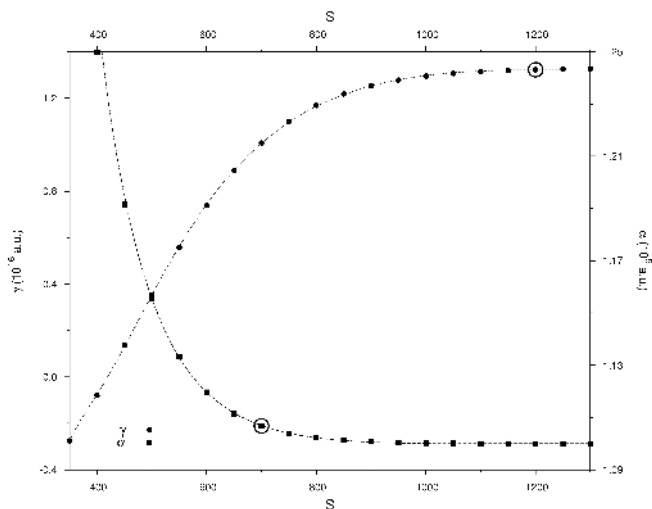


FIG. 7: PBE0 polarizability α_{xx} and second hyperpolarizability γ_{xxxx} of PA as functions of the shrinking factor (see text for discussion). Larger circles mark regions of the plot where α and γ values are converged to better than 1 % with respect to the asymptotes.

$\mathbb{E}^{\mathbf{k},vw}$ is the second derivative of the Lagrange multiplier matrix with respect to the field. We have used the symbols $\mathbb{F}^{\mathbf{k},u}$ and $\mathbb{F}^{\mathbf{k},uv}$ to indicate derivatives of the Fock matrix. The first derivative contains contributions from Ω as well as the two-electron terms, but only the latter contribute to the second derivative.

Details about the CPHF/KS method, and its implementation in the CRYSTAL program, can be found elsewhere;^{17–19} examples of application to materials of different dimensions are also available.^{75–79}

The all-trans isomer of polyacetylene (PA) is a good case to check the numerical accuracy of our implementation. PA is a prototype conjugated polymer semiconductor with π -electrons delocalized over a backbone of carbon atoms connected through alternating double and single bonds, with respective C-C bond lengths of 1.36 and 1.44 Å⁸⁰ (point symmetry C_{2h}). According to experiment,⁸¹ PA has a band gap E_g of 1.35 eV. Such a small band gap represents a severe test for DFT calculations of optical properties (especially pure DFT, which seriously underestimates the gap) because of inherent numerical instability as one approaches the conduction limit. The sensitivity with respect to the description of the electronic structure leads to calculated (hyper)polarizabilities that change by orders of magnitude depending on the approximation adopted. Finally, for finite oligomers it is known that conventional functionals strongly overshoot the correct values.

We computed the static longitudinal polarizability α_L and second hyperpolarizability γ_L of PA (the first hyperpolarizability β_L is null by symmetry) for Hartree-Fock (HF) and various density functionals.⁷⁷ The lowest estimates - *i.e.* $\alpha_L^{HF} = 165.2$ and $\gamma_L^{HF} = 6.16 \cdot 10^6$ a.u. - are provided by HF, which seriously overshoots

the energy gap ($E_g^{HF} = 6.8$ eV). Convergence of the calculated optical properties to the infinite periodic polymer limit is essentially achieved for a chain length of 50 monomers. This is consistent with an extensive literature on short PA oligomers, which also shows that HF values are relatively close to accurate Coupled-Cluster and Möller-Plesset results (see Ref. 82 for instance).

Pure density functionals shrink the energy gap of PA to less than 0.1 eV, thus leading to a catastrophic overshoot of the calculated optical properties. For the local density approximation (LDA),^{83,84} the lowest electronic transitions occur in the near Infra-Red ($E_g^{LDA} = 0.08$ eV) and the calculated optical properties soar to $\alpha_L^{LDA} = 1.10 \cdot 10^5$ and $\gamma_L^{LDA} = 1.32 \cdot 10^{16}$ a.u., respectively. As expected, hybrid functionals provide a more accurate description of the PA band structure. PBE0,⁸⁵ for instance, yields an optical gap equal to 1.43 eV. The corresponding (hyper)polarizabilities undergo a significant reduction compared to LDA, but still far exceed the HF values: $\alpha_L^{PBE0}/\alpha_L^{HF} \simeq 5$ and $\gamma_L^{PBE0}/\gamma_L^{HF} \approx 500$.

For small bandgap polymers, the precision of coupled perturbed (hyper)polarizability calculations is mainly determined by two computational parameters, namely

1. the shrinking factor used to generate a commensurate grid of \mathbf{k} points in the reciprocal space, and
2. the thresholds on the truncation of the two-electron integral series (see the CRYSTAL user's manual for details).

These two parameters are highly correlated because the set of lattice vectors employed to evaluate matrices in real space must also map the \mathbf{k} points spanned by the summations in eqs. (11) - (13). The more the band gap narrows, the more these parameters must be tightened. Thus, on the one hand, increasing the number of \mathbf{k} points naturally improves the description of the electronic band structure at \mathbf{k} values where $E_g(\mathbf{k}) \rightarrow 0$ (*i.e.* where $\alpha, \beta, \gamma \rightarrow \infty$). On the other hand, a larger number of two-electron contributions is required to account for the concurrent spread of the density matrix range in direct space.⁸⁶ Indeed, a very accurate PBE0 calculation of α and γ required consideration of 100 \mathbf{k} points in the first Brillouin zone and exchange contributions up to a distance of 250 Å from cell $\mathbf{0}$ in direct space to achieve well-converged results (Figure 7).

As implied above, the behavior of PA is exceptional because of its very small band gap. In general, the calculation of β and γ for insulators and semiconductors can be done with commonly used parameter settings, for example, a shrinking factor of 6-8 and the default value of $T_x = 12$ are adequate.^{19,87}

B. Electric Field Frequency Dependence

The calculated dependence of the electronic linear polarizability on the electric field frequency ω (or wavelength λ) can be directly compared with experimental

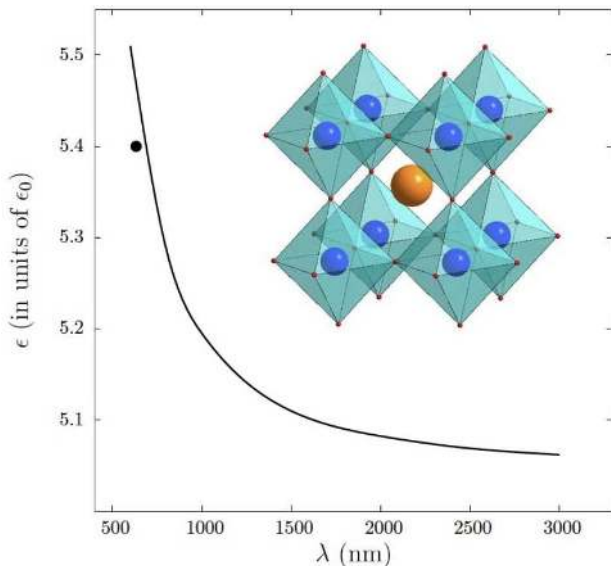


FIG. 8: Electronic contribution to the wavelength-dependent dielectric constant (continuous line), as computed with the PBE functional for the cubic phase of BaTiO₃. The experimental value, at $\lambda = 632.8$ nm, is reported as a black circle⁸⁸ and the structure of cubic BaTiO₃ is shown in the inset.

data in the high frequency limit. For the cubic phase of BaTiO₃ the dielectric constant was measured to be 5.41, in units of ϵ_0 , at $\lambda = 632.8$ nm.⁸⁸ We assume that this frequency is sufficiently high that the vibrational contributions are negligible. Then, in order to compare with experiment the frequency-dependent CPHF equations must be solved for the two first-order perturbation matrices, $U^{\mathbf{k},u}(+\omega)$ and $U^{\mathbf{k},u}(-\omega)$. Such calculations were carried out for the wavelength range 632.8 - 3000 nm.

Since, in our experience, the generalized-gradient approximation to DFT provides the best agreement with experimental dielectric tensors and photoelastic constants for most inorganic crystals (as compared to HF and hybrids),^{13–15,17,19} we used the PBE functional (Figure 8) for these calculations. It is clearly seen in Figure 8 that the variation of the electronic dielectric constant is quite large over the range of wavelengths considered, so that accurate calculation of ϵ at the actual wavelength of the experiment is important in this case and the agreement between the calculated and experimental values turns out to be within 2%.

C. Vibrational contributions to the (hyper)polarizabilities

Nuclear motions have a rôle in determining static and dynamic polarizabilities and hyperpolarizabilities. An efficient method for computing the vibrational contributions is the finite field nuclear relaxation (FF-NR) procedure proposed by Bishop, Hasan and Kirtman.⁸⁹ This method includes all terms that are harmonic or first-

order in electrical/mechanical anharmonicity. It is a general scheme that is currently implemented in CRYSTAL14 for non-periodic directions (3 independent directions in molecules, 2 in polymers, 1 in slabs) and for tensor components that couple periodic with non-periodic directions. The key step in the FF-NR method is a geometry optimization carried out in the presence of a finite, i.e. static, field. This optimization implicitly contains the information about harmonic and anharmonic vibrational parameters needed to obtain vibrational (hyper)polarizabilities. Thus, no force constants or electric property derivatives need to be explicitly calculated. If we denote the equilibrium geometry in a static electric field (ϵ) by R_ϵ , and R_0 without the field, then a Taylor series expansion of the field-dependent (electronic) dipole moment at the two geometries yields:

$$\begin{aligned} \mu_t(R_0) &= \mu_t(0, R_0) + \sum_u \alpha_{tu}^e \epsilon_u + \\ &+ \frac{1}{2} \sum_{u,v} \beta_{tuv}^e \epsilon_u \epsilon_v + \frac{1}{6} \sum_{u,v,w} \gamma_{tuvw}^e \epsilon_u \epsilon_v \epsilon_w + \dots \end{aligned} \quad (14)$$

and:

$$\begin{aligned} \mu_t(R_\epsilon) &= \mu_t(0, R_0) + \sum_u a_{tu}^\mu \epsilon_u + \\ &+ \frac{1}{2} \sum_{u,v} b_{tuv}^\mu \epsilon_u \epsilon_v + \frac{1}{6} \sum_{u,v,w} g_{tuvw}^\mu \epsilon_u \epsilon_v \epsilon_w + \dots \end{aligned} \quad (15)$$

In Eq. (14) the superscript “e” refers to the static electronic value (experimentally obtained by extrapolating to the static limit from measurements at sufficiently high frequency that vibrational contributions are negligible). In Eq.(15) the coefficients a^μ , b^μ and g^μ contain additional vibrational contributions that are determined by the geometry relaxation induced by the applied static field ϵ .⁸⁹

$$a_{tu}^\mu = \alpha_{tu}^e + \alpha_{tu}^{nr}(0; 0) \quad (16)$$

$$b_{tuv}^\mu = \beta_{tuv}^e + \beta_{tuv}^{nr}(0; 0, 0) \quad (17)$$

$$g_{tuvw}^\mu = \gamma_{tuvw}^e + \gamma_{tuvw}^{nr}(0; 0, 0, 0) \quad (18)$$

The static α^e and β^e also contain contributions due to anharmonic force constants and anharmonic electrical property derivatives. They give the vibrational contribution to $\beta^e(-\omega; \omega, 0)$, $\gamma^e(-\omega; \omega, 0, 0)$ and $\gamma^e(-2\omega; \omega, \omega, 0)$ in the high frequency limit (see Ref. 89). For the special case of the vibrational linear polarizability, one can also carry out a Berry phase treatment using the CRYSTAL code and we have found that the two methods give identical results.

As an example of the FF-NR method, calculations have been carried out on infinite periodic $(n,0)$ zigzag BN nanotubes.^{78,87,90} The largest tube considered ($n = 36$) contained 114 atoms in the unit cell. Our results

in that case show that the vibrational contributions to hyperpolarizabilities can exceed the electronic values as suggested by molecular calculations and theory. The vibrational contribution to each linear/nonlinear property increases with the nanotube radius. This increase is accompanied by an elliptical field-induced deformation of the cross-sectional geometry, which is enhanced for larger radius tubes due to their greater flexibility, along with a reduced band gap. The rate of increase relative to the corresponding static electronic property is dictated primarily by the number of static fields used to characterize the process. For the components considered it is larger in the transverse direction.

VI. VIBRATIONS IN SOLIDS: ANALYTICAL IR AND RAMAN INTENSITIES, VIBRATIONAL SPECTRA AND PHONON DISPERSION

CRYSTAL14 includes several new tools for characterizing the vibrational properties of crystalline solids. In particular, infrared (IR) and Raman spectra can be fully simulated, thanks to the newly introduced analytical computation of the peak intensities. This complements the calculation of peak positions available since CRYSTAL03. A procedure has been added to include Lorentz broadening and correction factors for the experimental setup. Thus, one can generate a spectrum that can be directly compared with experiment. Apart from the vibrations at the center of the Brillouin zone that are seen in IR and Raman spectra, CRYSTAL14 allows for the calculation of Phonon dispersion, a tool that is mandatory for investigation of thermodynamic properties.

A. Analytical Intensities

In previous releases of the code it was possible to compute the infrared intensity, but not the Raman intensity. The calculation was performed by means of numerical differentiation. A scheme involving Wannier functions was presented in CRYSTAL06 and a Berry phase scheme in CRYSTAL09. A new approach has been implemented for this release²²⁻²⁵ in which IR and Raman intensities are computed analytically and, as a result, very efficiently.

One possible way to obtain the working equations is to differentiate the expression for atomic gradients,^{91,92}

$$g_t^a = \frac{\partial E}{\partial r_{a,t}} = Tr \left[\frac{1}{2} (H + F)^{[a,t]} D - S_t^a D_w \right] \quad (19)$$

where the eigenvalue-weighted density matrix $D_w = \mathcal{F}[C^{\mathbf{k}\dagger} \varepsilon C^{\mathbf{k}}]$ has been introduced (\mathcal{F} here is the back-Fourier transform operator). H is the one-electron part of the Fock matrix F . The square brackets used for the superscript a, t in Eq. (19), indicate that differentiation is carried out only for integrals, and does not affect the density matrix. The aim of using Eq. (19) as a starting

point for our treatment here is to avoid explicit gradients of coefficients that would require the solution of an additional set of coupled-perturbed equations.

Born charges (Z^*) (and then the IR intensities) and the Raman tensor (I) can be obtained by differentiating Eq. 19 (at equilibrium geometry) one and two times, respectively, with respect to the electric field, and then imposing the zero field condition:

$$Z_{tu}^{a*} = \left. \frac{\partial g_t^a}{\partial \epsilon_u} \right|_{\epsilon=0} = Tr \left[\frac{1}{2} (\mathbb{H} + \mathbb{F})^{[a,t],u} D + \frac{1}{2} (H + F)^{[a,t]} \mathbb{D}^u - S \mathbb{D}_w^u \right]_{\epsilon=0} \quad (20)$$

$$I_{tuv}^a = \left. \frac{\partial^2 g_t^a}{\partial \epsilon_u \partial \epsilon_v} \right|_{\epsilon=0} = Tr \left[\frac{1}{2} (\mathbb{H} + \mathbb{F})^{[a,t],u,v} D + \frac{1}{2} \mathcal{P}_{u,v} \frac{1}{2} (\mathbb{H} + \mathbb{F})^{[a,t],u} \mathbb{D}^v + \frac{1}{2} (H + F)^{[a,t]} \mathbb{D}^{u,v} - S \mathbb{D}_w^{u,v} \right]_{\epsilon=0} \quad (21)$$

These equations can be recast in a form that is computationally more efficient and suitable for implementation. However, that requires some manipulations^{22,24} that go beyond the scope of this paper. We have also extensively discussed elsewhere^{22,25} that problems related to the geometric phase are entirely avoided.

The excellent numerical stability of the above procedure with respect to computational parameters such as reciprocal space sampling, integral screening thresholds and, to some extent, basis set has been demonstrated.²³ Despite its quite recent implementation, a number of applications have been carried out on interesting crystalline systems, like quartz,²³ spessartine,²⁴ pyrope,⁹³ jadeite,⁹⁴ UiO-66²³ and CPO-27⁹⁵ metal-organic frameworks. In the following sub-section we present a new study on vibrational properties of Forsterite.

B. Simulated Vibrational Spectra

The infrared oscillator strengths f_p can be computed for each p -th mode by means of the mass-weighted effective mode Born charge vectors \vec{Z}_p .^{97,98}

$$f_{p,tu} = \frac{1}{4\pi\epsilon_0} \frac{4\pi}{V} \frac{Z_{p,t} Z_{p,u}}{\nu_p^2}, \quad (22)$$

$$Z_{p,t} = \sum_{a,u} \mathbf{t}_{p,au} Z_{tu}^{a*} \frac{1}{\sqrt{M_a}}, \quad (23)$$

where ϵ_0 is the vacuum dielectric permittivity ($1/4\pi\epsilon_0 = 1$ atomic unit), V is the unit cell volume, t and u refer to the Cartesian components, $\mathbf{t}_{p,au}$ is an element of the eigenvectors matrix \mathbf{T} of the mass-weighted Hessian matrix \mathbf{W} , that transforms the Cartesian atomic directions into the p -th normal coordinate directions (see Section VIC below).

TABLE IV: IR properties of Mg_2SiO_4 forsterite computed at the B3LYP level: frequencies ν (in cm^{-1}) and oscillator strengths f (dimensionless).

B _{3u}			B _{2u}			B _{1u}		
#	ν	f	#	ν	f	#	ν	f
1	206.2	0.0248	14	143.1	0.0827	27	206.5	0.0020
2	274.9	0.0559	15	277.2	0.0753	28	277.6	0.2059
3	293.7	0.4393	16	292.2	1.6219	29	290.3	1.3639
4	322.2	0.0554	17	350.2	1.4693	30	313.0	0.0080
5	387.8	1.2456	18	403.3	0.2651	31	419.6	0.9781
6	411.6	1.2740	19	431.7	0.2937	32	427.9	0.2979
7	475.9	0.0128	20	464.8	0.1664	33	489.5	0.2399
8	513.4	0.3391	21	517.3	0.0444	34	513.4	0.4134
9	540.0	0.0048	22	534.5	0.2105	35	874.4	0.5953
10	613.7	0.2044	23	637.6	0.0002			
11	838.1	0.0053	24	835.1	0.1342			
12	961.9	0.2283	25	870.3	0.3548			
13	982.4	0.2415	26	988.9	0.0078			

A simulated reflectance curve $R_{tt}(\nu)$ along the tt direction can be obtained by means of the Fresnel formula:⁹⁹

$$R_{tt}(\nu) = \left| \frac{\sqrt{\epsilon_{tt}(\nu) - \sin^2(\theta)} - \cos(\theta)}{\sqrt{\epsilon_{tt}(\nu) - \sin^2(\theta)} + \cos(\theta)} \right|^2, \quad (24)$$

where θ is the incidence angle of the IR beam with respect to the normal to the surface and $\epsilon_{tt}(\nu) = \epsilon_{1,tt}(\nu) + i\epsilon_{2,tt}(\nu)$ is the tt -th component of the complex dielectric function. The maxima of $\epsilon_2(\nu)$ and of $\text{Im}(-1/\epsilon(\nu))$ (Loss Function) correspond to the TO and LO frequencies, respectively. Note that, when the symmetry of the system is orthorhombic or higher, $\epsilon(\nu)$ is a diagonal tensor, so that only the xx , yy and zz components are non-null. The classical Drude-Lorentz model⁹⁹ describes the dielectric function as a superposition of damped harmonic oscillators:

$$\epsilon_{tt}(\nu) = \epsilon_{\infty,tt} + \sum_p L_{p,tt}(\nu), \quad (25)$$

where $\epsilon_{\infty,tt}$ is the high-frequency (electronic) dielectric contribution and the oscillator $L_{p,tt}(\nu)$ is defined as:

$$L_{p,tt}(\nu) = \frac{f_{p,tt}\nu_p^2}{\nu_p^2 - \nu^2 - i\nu\gamma_p}. \quad (26)$$

Each oscillator is characterized by three parameters: the frequency ν_p of the TO mode (note: only of the TO mode), its strength along the tt direction $f_{p,tt}$ (related to the plasma frequency $\tilde{\nu}_p$ through $f_p = \tilde{\nu}_p^2/\nu_p^2$) and the damping factor γ_n . As the implemented harmonic model does not permit to compute values for the latter quantity, a guess or average value is usually taken²⁶.

TABLE V: Raman properties of Mg_2SiO_4 forsterite computed at the B3LYP level: frequencies ν (in cm^{-1}) and polarized intensities (arbitrary units; values are renormalized so that the highest one is set to 1000).

A _g				
#	ν	a ²	b ²	c ²
1	188.4	7	2	2
2	233.9	14	5	10
3	307.1	29	0	8
4	328.8	3	23	10
5	344.6	15	19	12
6	424.8	14	7	2
7	559.6	5	141	136
8	618.2	328	0	2
9	819.2	519	850	187
10	856.1	1000	323	839
11	967.3	188	125	172

B _{1g}			B _{2g}			B _{3g}		
#	ν	d ²	#	ν	e ²	#	ν	f ²
12	224.9	1	23	182.7	2	30	190.0	2
13	260.4	2	24	252.9	3	31	303.6	4
14	317.4	24	25	323.6	2	32	322.3	2
15	367.1	1	26	374.0	14	33	380.5	42
16	391.5	9	27	450.7	57	34	420.8	17
17	441.9	57	28	608.3	58	35	608.9	143
18	596.2	44	29	884.4	85	36	928.0	129
19	644.9	10						
20	834.8	32						
21	865.5	39						
22	979.7	6						

When simulating the experimental Raman spectrum of a real crystal, a number of factors must be taken into account. The relevant formulas, which are well-known, are briefly summarized here for ease of reference. For an oriented single-crystal the Raman Stokes scattering intensity associated with the general tu component of the polarizability tensor corresponding to the p -th vibrational mode of frequency ω_p may be calculated as:

$$I_{tu}^p \propto C \left(\frac{\alpha_{tu}}{\partial Q_p} \right)^2 \quad (27)$$

where Q_p is the the normal mode coordinate for mode p . The prefactor C depends¹⁰⁰ on the laser frequency ω_L and the temperature T :

$$C \sim (\omega_L - \omega_p)^4 \frac{1 + n(\omega_p)}{30\omega_p} \quad (28)$$

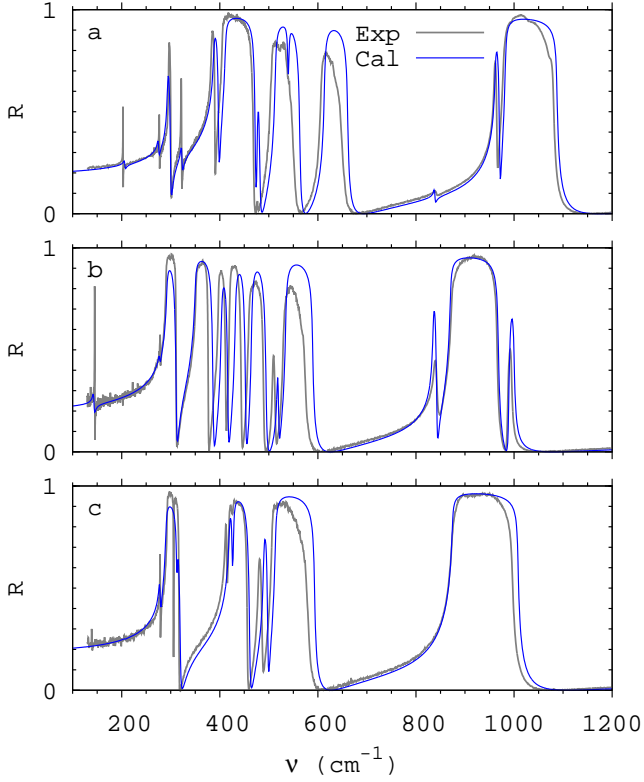


FIG. 9: IR reflectance spectra of Mg_2SiO_4 forsterite along the a , b and c axes (B_{3u} , B_{2u} and B_{1u} symmetries, respectively) computed at the B3LYP level. The damping factor γ is set to 3 cm^{-1} for all modes. Experimental spectra from Ref. 96 are reported for sake of comparison.

with the Bose occupancy factor $n(\omega_p)$ given by

$$1 + n(\omega_p) = \left[1 - \exp\left(-\frac{\hbar\omega_p}{k_B T}\right) \right]^{-1} \quad (29)$$

The polycrystalline (powder) spectrum can be computed by averaging over the possible orientations of the crystallites as described in Eq. (4) and (5) of Ref. 101.

While the intensity of the transverse optical (TO) modes is straightforwardly computed once the appropriate polarizability derivative is obtained, the corresponding calculation for longitudinal optical (LO) modes requires a correction due to $\chi_{uvw}^{(2)}$.^{100,102}

$$\left. \frac{\partial \alpha_{uv}}{\partial r_{a,t}} \right|_{\mathcal{R}_0} = \frac{1}{V} \left. \frac{\partial^3 E^{TOT}}{\partial r_{a,t} \partial \mathcal{E}_u \partial \mathcal{E}_v} \right|_{\mathcal{E}=0, \mathcal{R}_0} - 2 \sum_{u'} Z_{u't}^* \sum_w \epsilon_{u'w}^{-1} \chi_{uvw}^{(2)} \quad (30)$$

In Eq. (30) ϵ^{-1} is the inverse of the high-frequency (*i.e.* pure electronic) dielectric tensor. $\chi^{(2)}$ is defined as in Eq. (69) of Ref. 17.

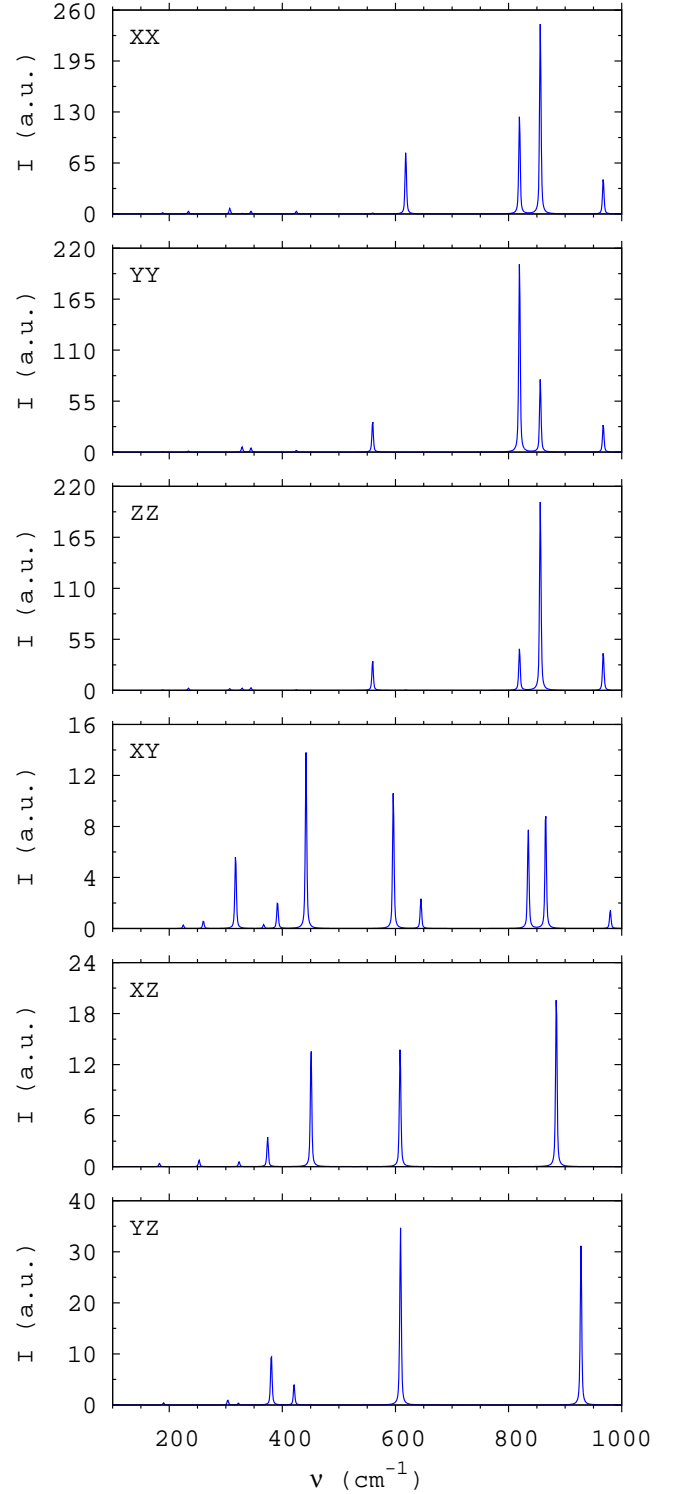


FIG. 10: Polarized Raman spectra of Mg_2SiO_4 forsterite computed at the B3LYP level with polarizations: XX (a^2 contributions), YY (b^2), ZZ (c^2), XY (d^2), XZ (e^2), YZ (f^2). The damping factor γ is set to 3 cm^{-1} for all modes.

1. IR and Raman spectra of forsterite

Mg_2SiO_4 forsterite, is the magnesium end member of the olivine family, that are important rock-forming silicates. As a demonstration of the methods presented in this section, we show here their application to the simulation of the vibrational properties of this crystal.

In Figure 9 the simulated reflectance spectrum along the three axes a, b, c is compared with experiment. The similarity of the two is striking, with all the main features of the spectrum being correctly reproduced by the simulation. Note in particular that the oscillator strength affects the width of the reflectance bands and peaks (the larger the value, the wider the band); experimental wide bands and narrow peaks always correspond to computed features of the same kind. Table IV reports the main ingredients required for the simulated spectra: the IR-active TO vibrational frequencies and the corresponding oscillator strengths; one more ingredient (not reported) is the dielectric tensor computed by means of CPHF module of CRYSTAL14.

In Figure 10 the computed Raman spectra are reported for the six independent orientations. It is interesting to notice how, for this crystal, these directions indeed provide significantly different spectra. This information is very useful for experimentalists in order to obtain a perfect orientation of the crystal and thus avoid leakage from one symmetry to another. We plan in the future to compare these data with high-quality measured spectra. Table V reports the raw computed data that were used to generate Figure 10. Note how the Raman intensities range over more than four orders of magnitude, and that several modes that are considered Raman-active upon symmetry analysis are found to possess zero or very low intensity. Another interesting aspect (that is not seen in Figure 10 due to the adopted scale) is that B_{1g} , B_{2g} and B_{3g} symmetries have considerably lower absolute intensities compared to the A_g modes. This highlights the potential interest in measuring accurate directional spectra since these modes (specially the B_{1g} ones) would be hardly seen in a polycrystalline (powder) spectrum.

C. Phonon Dispersion and Thermodynamic properties

The calculation of vibration frequencies at the Γ point ($\mathbf{k} = \mathbf{0}$, at the center of the First Brillouin Zone -FBZ- in reciprocal space), within the harmonic approximation, is available from the CRYSTAL03 version of the program.^{20,21} The vibration frequencies at the center of the FBZ (directly comparable with the outcomes of IR and RAMAN measurements), are obtained from the diagonalization of the mass-weighted Hessian matrix of the second derivatives of the total energy per cell with respect

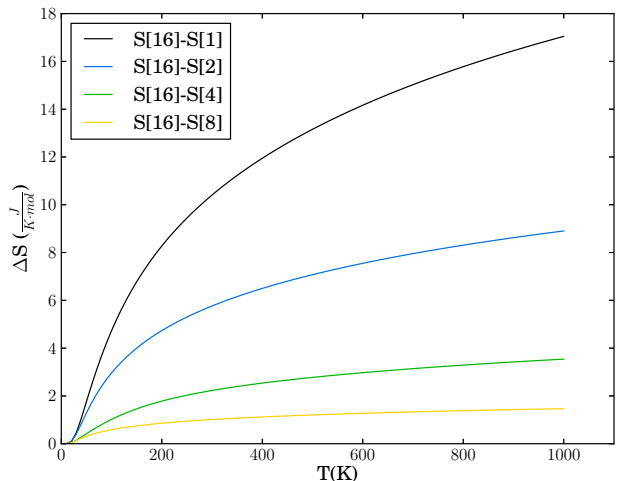


FIG. 11: Differences of the entropy S of pyrope, as a function of temperature, as computed with different supercells with respect to the largest x16 supercell considered. Data obtained with the B3LYP hybrid functional.

to atomic displacements u :

$$W_{at,bu}^{\Gamma} = \frac{H_{at,bu}^0}{\sqrt{M_a M_b}} \quad \text{with} \quad H_{at,bu}^0 = \left(\frac{\partial^2 E}{\partial u_{at}^0 \partial u_{bu}^0} \right), \quad (31)$$

where atoms a and b (with atomic masses M_a and M_b) in the reference cell are displaced along the t -th and u -th Cartesian directions. The first derivatives of the total energy per cell ($g_t^a = \partial E / \partial u_{at}$) with respect to atomic displacements from the equilibrium configuration \mathcal{R}^{eq} are computed analytically, whereas second derivatives numerically, using a two-point formula:

$$\frac{\partial^2 E}{\partial u_{at} \partial u_{bu}} \approx \frac{g_t^a(\mathcal{R}^{\text{eq}}, u_{bu} = +\bar{u}) - g_t^a(\mathcal{R}^{\text{eq}}, u_{bu} = -\bar{u})}{2\bar{u}},$$

where $\bar{u} = 0.003 \text{ \AA}$, a value 10 - 50 times smaller than that usually used in other solid state programs.¹⁰³⁻¹⁰⁵

The calculation of the thermodynamic properties is more demanding, as it requires the knowledge of phonon modes over the complete FBZ; phonons at points different from Γ can be obtained by considering a supercell (SC) of the original unit cell, following the so called *direct method*.^{106,107} The lattice vectors $\mathbf{g} = \sum_t l_t^g \mathbf{a}_t$ identify the general crystal cell where $\{\mathbf{a}_t\}$ are the direct lattice basis vectors, with $t = 1, \dots, D$ (where D is the dimensionality of the system: 1, 2, 3 for 1D, 2D, 3D periodic systems): within Born von Kármán periodic boundary conditions the integers l_t^g run from 0 to $L_t - 1$. The parameters $\{L_t\}$ define the size and shape of the SC in direct space. Let us label with \mathbf{G} the general super-lattice (*i.e.* whose reference cell is the SC) vector and let us introduce the $L = \prod_t L_t$ Hessian matrices $\{\mathbf{H}^{\mathbf{g}}\}$ whose elements are $H_{at,bu}^{\mathbf{g}} = \partial^2 E / (\partial u_{at}^0 \partial u_{bu}^0)$ where, at variance with equation (31), atom b is displaced in cell \mathbf{g} , along with all its periodic images in the crystal, in cells $\mathbf{g} + \mathbf{G}$.

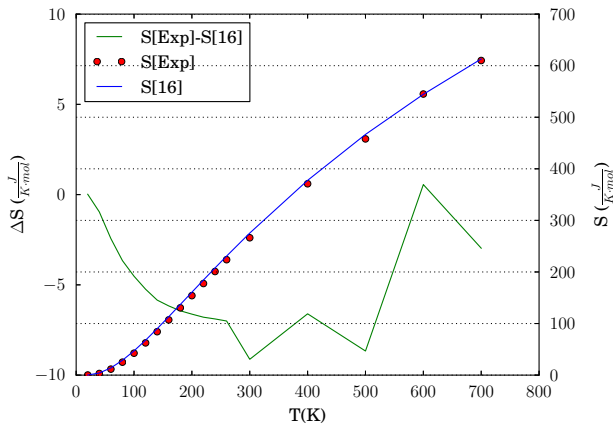


FIG. 12: Calculated and experimental^{109–111} entropy for pyrope and difference between the two, as a function of temperature. Data obtained with the B3LYP hybrid functional.

The set of L Hessian matrices $\{\mathbf{H}^{\mathbf{g}}\}$ can be Fourier transformed into a set of *dynamical matrices* $\{\mathbf{W}^{\mathbf{k}}\}$ each one associated with a wavevector $\mathbf{k} = \sum_t (\kappa_t/L_t) \mathbf{b}_t$ where $\{\mathbf{b}_t\}$ are the reciprocal lattice vectors and the integers κ_t run from 0 to $L_t - 1$:

$$W_{at,bu}^{\mathbf{k}} = \sum_{\mathbf{g} \in SC} \frac{H_{at,bu}^{\mathbf{g}}}{\sqrt{M_a M_b}} \exp(i\mathbf{k} \cdot \mathbf{g}), \quad (32)$$

The eigenvalues of the dynamical matrices are the square of the vibrational frequencies $\nu_p^{\mathbf{k}}$, while the eigenvectors correspond to the normal modes. The frequencies $\nu_p^{\mathbf{k}}$ define the energy spectrum of the harmonic oscillators:

$$E_p^{\mathbf{k}}(n) = h\nu_p^{\mathbf{k}} \left(\frac{1}{2} + n \right). \quad (33)$$

The crucial point is then the speed of convergence of the integration in reciprocal space, that is substituted by a finite sum over the \mathbf{k} points, a large number of points implying calculations with a large supercell. Thermodynamic properties of $\text{Mg}_3\text{Al}_2(\text{SiO}_4)_3$ pyrope garnet are currently under investigation.¹⁰⁸ Figure 11 shows the convergence of entropy as a function of the supercell size, from the 80 atoms primitive cell, indicated as x1, up to x16, a supercell containing $16 \times 80 = 1280$ atoms (thus permitting to obtain the frequencies at 16 \mathbf{k} points. All the considered supercells do have the cubic symmetry, so as to exploit it. In all cases, only 9 SCF+G calculations are performed. The figure shows, for various temperatures, the difference for the various supercells with respect to the x16 data. It turns out that already at x8 entropy is quite well converged, the difference with respect to x16 being smaller than 1 kJ/mol. In Figure 12 our *ab initio* results are compared with experimental data.^{109,110} The overall agreement is quite good, much better than obtained previously.¹¹¹

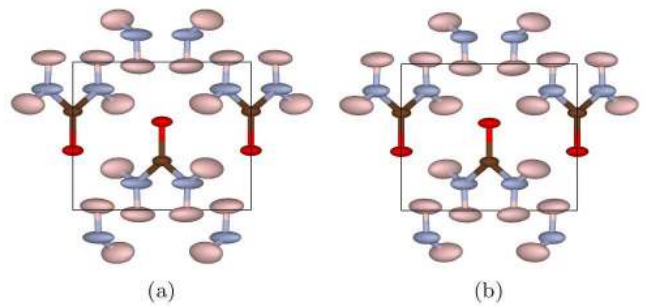


FIG. 13: Urea ADPs at 123 K as (a) computed on a $2 \times 2 \times 2$ supercell³⁰ and (b) obtained experimentally.¹¹³ Ellipsoids are drawn at 75% probability.

1. Anisotropic Displacement Parameters

Another important consequence of thermal motion is the damping of X-ray and Neutron Scattering diffracted intensities as a function of temperature (see Section VIII B). These effects are usually accounted for by means of the mean square displacements (MSD) along each normal mode of the harmonic oscillators. The sum of these MSD over all the oscillators leads to the definition of the Anisotropic Displacement Parameters (ADPs).¹¹² These quantities can now be computed automatically with the CRYSTAL14 version of the program.²⁹ ADPs have recently been computed with such a scheme for several molecular crystals.³⁰ Good agreement was found with respect to experimental data. ADPs convergence was studied with respect to several computational variables, such as basis set, lattice parameters, and Hamiltonian. In Figure 13 we report a graphical representation of the computed and experimental¹¹³ ADPs of crystalline urea. The overall features of atomic thermal motion are well described, with only the computed ADPs of nitrogen overshooting the experimental ones.

The effect of thermal motion on the electronic charge density (ECD) and structure factors is treated in Section VIII B, where the case of crystalline silicon is considered.

VII. TOOLS FOR STUDYING SOLID SOLUTIONS AND DISORDERED SYSTEMS

Solid solutions and disordered materials are characterized by non periodic occupation of some sites. Experimental structure determinations normally interpret the site-occupancy pattern in terms of average fractional site occupations. The experimental space group of disordered crystals or solid solutions cannot be used in the simulation of such systems, because fractional occupancies of crystallographic sites cannot be adopted as such. Each average occupation corresponds to a number of cell configurations (*i.e.* distributions of the atoms or vacancies involved in these fractionally occupied sites) from the collection of which the observed crystal properties

arise. Evaluating such average properties with quantum-mechanical simulations is a tremendous challenge due to the convolution of two main requirements. First, in order to account for non periodic site occupation and to keep periodicity, multiple cells - supercells ($n_a \mathbf{a} \times n_b \mathbf{b} \times n_c \mathbf{c}$) - have to be considered, where n_a , n_b , n_c are integers and \mathbf{a} , \mathbf{b} and \mathbf{c} are the vectors defining the primitive experimental cell. As a consequence, structure relaxation becomes very expensive. Second, the number of configurations increases as R^D , where R and D are the number of considered species and sites, respectively. The extreme difficulty of such a task has seriously hindered first-principle quantum-mechanical approaches so far.

In order to reduce the computational cost, the action of symmetry on the configurations can be considered. Before presenting the tools dedicated to the study of disordered systems and implemented in the new version of CRYSTAL, few concepts about the action of symmetry will be very briefly sketched. For full details, the reader can refer to Ref. 31 and 32.

A. Brief overview on the action of symmetry

Consider a structure (of any dimensionality) characterized by a symmetry group G and possessing an irreducible crystallographic position d of multiplicity $|D|$, occupied by atomic species A. Using Wyckoff's notation, this position would be noted $|D|d$. This structure is known as the aristotype for the considered disordered system.

Suppose atomic species X substitutes for A in any proportion on d site. Then, $|D| + 1$ compositions are possible : $A|D| - \alpha X$, $\alpha = 0, \dots, |D|$. If positions of A and X cannot be distinguished, the compound is said disordered. For each composition, there are

$$|S_\alpha| = \binom{|D|}{\alpha} = \frac{|D|!}{\alpha!(|D| - \alpha)!}$$

possibilities to place atoms A and X. Summing over the $|D| + 1$ composition, $|S| = 2^{|D|}$ possibilities exists. Each possibility is called a configuration. In Figure 14, two-color configurations derived from a C_{4v} aristotype structure are presented as a function of the composition. The various species are sketched by colored circles. The composition is expressed as the ratio of the different species. For more than two species, the number of configurations can easily be calculated. In the present version of CRYSTAL, 2 species only are allowed.

The aristotype symmetry group naturally partitions the set of configurations (S) into subsets of symmetry related configurations. Two configurations are symmetry related or equivalent if it exists an operator $g \in G$ that maps one onto the other. No element of G relates configurations belonging to two subsets. So, subsets are symmetry-independent classes of configurations (SICs). Properties of configurations belonging to the same SIC

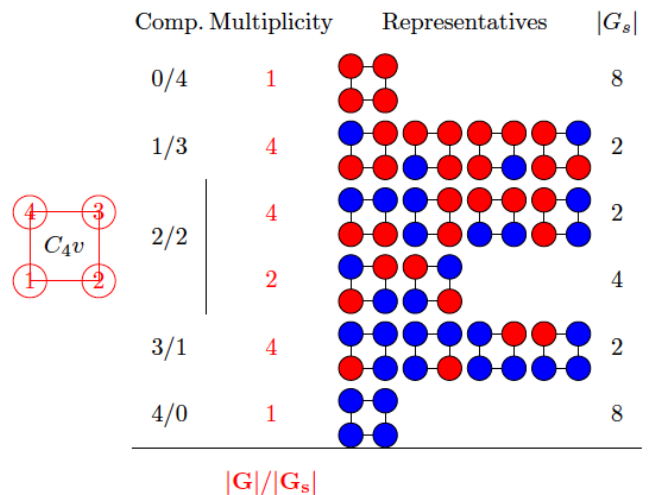


FIG. 14: (color online) Symmetry-independent classes of two-color configurations on four positions resulting from the action of the C_{4v} group. \mathcal{M} is the number of configurations per class (*multiplicity*); $|G_s|$ is the number of residual symmetry operators in each class. The multiplicity of each class equals $|G|/|G_s|$.

being identical or equivalent, each SIC can be fully described by one representative. Any element of the class can be chosen as a representative. The contribution of a given SIC to average properties depends on its multiplicity (\mathcal{M}) or the number of configurations it contains, modulated by Boltzmann distribution. The multiplicity of a given SIC is related to the symmetry of its elements $\mathcal{M} = |G|/|G_s|$ where $|G_s|$ is the group of any chosen configuration in the SIC. Multiplicity and $|G_s|$ are given in Figure 14. Obviously, two configurations belonging to the same SIC have the same composition, but two configurations sharing the same composition may belong to different classes (see again Figure 14, composition 2/2).

So, the number of SICs is a key quantity. Counting the number of SICs relies on Polya's theory that exploits the Cauchy-Frobenius lemma, often called the Burnside lemma:

$$|\Delta(S)| = \frac{1}{|G|} \sum_{g \in G} |S_g|,$$

where $\Delta(S)$, $|\Delta(S)|$ and S_g are the set of SICs, the number of SICs in S and the set of configurations stabilized by g or configurations whose symmetry group contains g , respectively. The $|S_g|$ are easily estimated using the Cayley theorem that states that any group acting on a set of $|D|$ points is isomorphous of a subgroup of the group of permutations of $|D|$ objects. As such, the symmetry operators identified to permutations can be represented by cycles. Considering the example in Figure 14, d_1 's action about the plane passing on 1 and 3 is described by (1)(3)(24) while (1234) accounts for the action of the 4-fold rotation. It follows that a configuration is stabilized by d_1 if positions in a given cycle are occupied by the

TABLE VI: Total number of SICs N^D resulting from the action of the space group $Ia\bar{3}d$ on sets of sites (*Dod* : dodecahedral ; *Oct*: octahedral) of the garnet structure considering two species. The primitive cell contains 8 dodecahedral and 12 octahedral sites. The conventional cell is twice.

	N^{Dod}	N^{Oct}
Primitive cell	23	154
Conventional cell	874	179'444

same species. By induction, if $|Cyc_D(g)|$ is the number of cycles for operation g acting on D , $|S_g| = |R|^{|Cyc_D(g)|}$. The number of SICs is given by Polya's formula:

$$|\Delta(S)| = \frac{1}{|G|} \sum_{g \in G} |R|^{|Cyc_D(g)|}. \quad (34)$$

Polya's formula is not a tool to generate the configurations and their classes. In the 2-species case, efficient algorithm relies on lexicographic representation and ordering of the configurations briefly described in the next section. Simple applications show that the number of SICs increases very quickly with the number of involved sites and the number of species. In Table VI, applications to the garnet structure considering solid solution on the dodecahedral or the tetrahedral sites illustrate this point. Despite factorization by the number of SICs, *ab initio* calculations might remain unfeasible for a large number of situations. For large systems, it becomes impossible to generate the full list of SICs and to compute quantum mechanically all of them. Some authors have proposed to compute randomly selected configurations. However, such an approach suffers drawbacks. For example, equivalent calculations might be performed. The SICs are found as a function of their multiplicity, the probability to reach a given SIC of multiplicity \mathcal{M} equals $\mathcal{M}/|S|$. So, the longer (low symmetry) the SIC, the larger its probability to be found. As the size of the cell increases, the lower the probability to reach symmetric SICs. However, several lines of thinking suggest that the most stable configuration should have high symmetry. If so, it is important to reach the SICs independently of their symmetry.

This can be obtained by rewriting Polya's formula. Two elements g and g' lying in the same conjugacy class can be exchanged, so they have the same cycle structure : $|Cyc_D(g)| = |Cyc_D(g')|$. Then, Eq. 34 can be factorized by conjugacy classes,

$$|\Delta(S)| = \frac{1}{|G|} \sum_{j=1}^{|C|} |\mathcal{C}_j| |R|^{|Cyc_D(g_j)|},$$

where g_j is a representative of the class \mathcal{C}_j . From this expression, probability distribution on the set of conjugacy

classes: $\mathcal{C} = \{\mathcal{C}_1, \dots, \mathcal{C}_{|C|}\}$ can be defined:

$$\sum_{j=1}^{|C|} \frac{|\mathcal{C}_j| |R|^{|Cyc_D(g_j)|}}{|\Delta(S)| |G|} = 1, \quad (35)$$

with

$$\mathbf{Prob}(\mathcal{C}_j) = \frac{|\mathcal{C}_j| |R|^{|Cyc_D(g_j)|}}{|\Delta(S)| |G|}, \quad j = 1, \dots, |C|. \quad (36)$$

For G and D given, the probability of the conjugacy classes can be calculated. If a configuration s stabilized by an operator g belonging to a given class is randomly constructed, the probability that this configuration is in a given SIC ω is $\mathbf{Prob}(\omega \ni s | s \in S_g) = |\omega \cap S_g| / |S_g|$. This probability is different from the previous one. It can be shown that selecting the conjugacy classes with such probability (36) and building at random a stabilized configuration for each selected class, SICs are found with an equal probability $1/|\Delta(S)|$. A configuration stabilized by an operation of symmetry is obtained mapping all elements of the same cycle onto the same species. As a consequence, only the identity class allows to produce asymmetric configuration because every cycle contains one element. If the probability of the identity is set to zero, only symmetric configurations show up. The probability of the SICs becomes a decreasing function of the multiplicity. The most symmetric SICs have the highest probability to be found.

B. Implemented tools

Tools dedicated to the study of disordered materials offer the possibility to count and generate lists of representatives as well as the multiplicity of SICs. Disordered systems involving several symmetry independent sites can be considered, but only isoelectric substitutions should be analyzed.

Counting and enumerating SICs

The CONFCONT keyword yields in lexicographic order representatives of the SICs. The lexicographic order is equivalent to the alphabetical order. The $|D|$ involved sites are arbitrarily ordered from 1 to $|D|$, the two species are identified as 0 and 1 (0 foregoes 1). Each configuration is represented by a "string" (L) of length $|D|$ composed by 0 and 1. Then a configuration s_1 foregoes a configuration s_2 if for $1 \leq i \leq |D|$, $L_1(i) < L_2(i)$, $L_1(j < i) = L_2(j < i)$. For example, $|D| = 4$, (0010) foregoes (0011) which foregoes (0100). This natural order permits to perform efficient tests to identify new SICs and their representative without requiring memory storage.

Default options ($|D| + 1$ compositions, symbol for the substituting species, multiplicity, etc.) can be modified by means of specific keywords. Among those, ONLYCOMP

permits to select a specific composition, thus reducing the length of the output. The number of SICs over the $|D| + 1$ compositions can be obtained without generating the SICs at no computer cost. We warn the user to use this option when starting a new study in order to avoid lengthy output.

Two-body interactions up to a chosen distance are symmetry-sorted invoking the `INTPRT` keyword. Higher-order interactions are not considered.

Random selection of SICs

The `CONFRAND` keyword calls a symmetry-adapted and tunable Monte Carlo (not Metropolis) sampling tool. At a given composition, it returns representatives of SICs and their multiplicity. By default, SICs are found with the same probability, but symmetric SICs can be selected. The number of SICs for the given composition is calculated. The number of searched SICs might be user-controlled. The number of Monte Carlo draws can be chosen, in this case the number of found SICs is not predefined.

Calculations on selected configurations

The `CONFRAND` option produces a file that allows to perform calculations on found SICs. These calculations are launched by the `RUNCONF` keyword. Taking advantage of the parallel architecture, a multitask scheme is implemented, so several configurations can be calculated simultaneously, each one in parallel mode. The transfer file is quite compact because each configuration is simply identified by an integer that carries full informations about distribution of the species on the different sites.

Results on normal spinel MgFeAlO_4 are shown in Figure 15. The conventional cell has been considered. Mixing occurs on the 16 octahedral sites. The average energy has been calculated producing SICs according to two schemes. In one scheme, the SICs have found uniformly at random independently of their symmetry. The second one contains two steps, symmetric SICs are first searched, then the asymmetric ones. The running average is compared to the limit value (dashed horizontal line). In both cases, the convergence is quite rapid. But at low temperature, searching the symmetric SICs first seems more efficient supporting the idea that the stablest configuration has some high symmetry.

Defining the supercell

As previously stated, simulation of disordered systems requires the use of supercells. In such cases, full point and translational symmetries of the supercell have to be properly considered in order to generate a minimal list of SICs. This includes non conventional centering of the

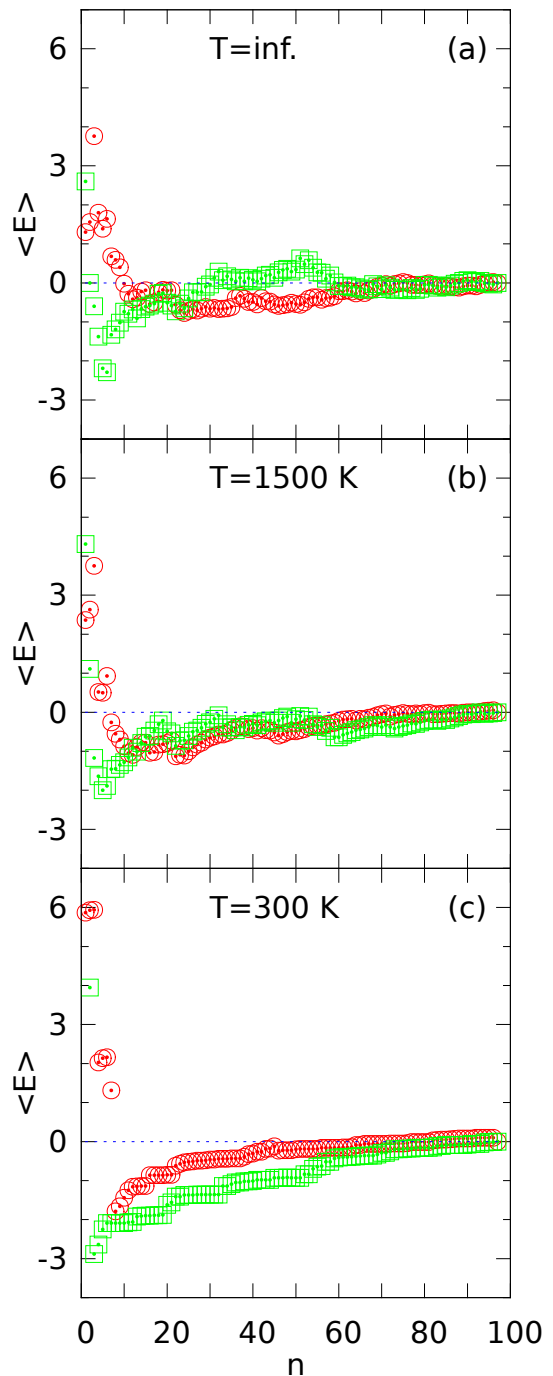


FIG. 15: (color online) Spinel case study: deviation of the Boltzmann average energy E from its limit value over the full set of SICs, in mHartree, as a function of the number of considered configurations n . Panels (a), (b) and (c) correspond to the infinite temperature limit, 1500 K and 300 K, respectively. For each temperature, the Boltzmann average energy has been calculated according to two procedures over the full set of configurations. Circle: MC sampling, square: symmetry-enhanced sampling.

supercells. Starting from a primitive cell, suppose that a $n_a \mathbf{a} + n_b \mathbf{b} + n_c \mathbf{c}$ supercell is built. This supercell contains

$n_a \times n_b \times n_c$ translation vectors corresponding to symmetry operations. Then, the space group of the supercell to be considered is a non-conventional one and includes the operators of the group of the aristotype compatible with the supercell combined with the $n_a \times n_b \times n_c$ translational operators. A special supercell option (SCELCONF keyword) allows to deal with this non-conventional space groups.

VIII. ELECTRON DENSITY ANALYSIS

Detailed information about the electronic structure of crystalline compounds is provided by observables related to the one-electron density matrix (DM), such as the electron charge density (ECD) $[\rho(\mathbf{r})]$, and the electron momentum density (EMD) $[\pi(\mathbf{p})]$.¹¹⁴ The ECD is obtainable from diffraction experiments and is straightforwardly related to the topological features of the system in direct space, thus to position of nuclei and characteristics of bonds. The EMD can be reconstructed from directional Compton scattering experiments:¹¹⁵ the analysis of the distribution in momentum space of the slow valence electrons is known to provide valuable complementary insight into the chemical features of the system.

Some of the new features of CRYSTAL14 as regards electron densities are: i) the complete topological analysis of the ECD by means of the automated integration of the TOPOND package into CRYSTAL14 (see Section VIII A); ii) the parallelization, with linear speed-up, of all the algorithms related to ECD and EMD; iii) calculation of anisotropic displacement parameters and Debye-Waller thermal factors for dynamical X-ray structure factors (see Section VIII B); iv) new algorithms for the analysis of the EMD (see Section VIII C).

A. Topological analysis of Charge Density

Over the past decade, studies of chemical bonding in solids have experienced a renovated interest. Among other reasons, the availability of reconstructed experimental electron densities, derived from high resolution synchrotron radiation diffraction measurements,^{116,117} and accurate *ab initio* theoretical determinations, has provided a unique opportunity for comparison, mutual validation and enhancement in the analytical skills of both approaches.^{117,118}

The Quantum Theory of Atoms in Molecules¹¹⁹ (QTAIM), as implemented within the theoretical framework of the CRYSTAL program¹²⁰ (QTAIMAC) represents the most complete density-based topological tool for chemical bonding studies and, since 1998, has been implemented by C. Gatti in the public code TOPOND.³³ Currently, TOPOND has been embedded into the CRYSTAL14 suite of programs and the whole machinery is now easily accessible through the keyword TOPO of its PROPERTIES module. The QTAIMAC scheme explicitly refers

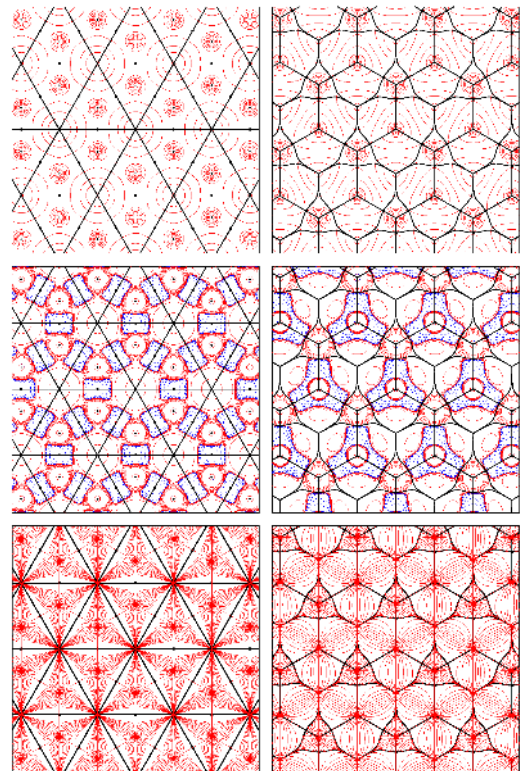


FIG. 16: From top to bottom: electron density, $\rho(\mathbf{r})$, its Laplacian, $\nabla^2\rho(\mathbf{r})$ and the $\nabla\rho(\mathbf{r})$ trajectories for graphene (left panels) and boron nitride (right panels). Bond paths are indicated in heavy black lines and the bond critical points are denoted by filled circles. Dotted blue lines indicate negative contour levels.

to the experimentally observable electron density and yet involves different quantities whose analysis can be carried out separately but whose properties have to be combined and discussed together in order to get a complete and reliable picture of the underlying bonding network.

A first step of the analysis provides so-called critical points (CP) of space, that are points where the gradient of any scalar function vanishes. It is worth noting that a two-way correspondence between CPs of the density, $\nabla\rho(\mathbf{r}) = 0$, and chemically recognizable structures such as nuclei, bonds, rings and cages, is always possible.¹²⁰ Then, the topological analysis of the Laplacian of the electron density, $\nabla^2\rho(\mathbf{r})$, can reveal the atomic shell structure and the degree of sharing of paired electrons among neighboring atoms. In particular, it has been shown that local maxima and minima in $\nabla^2\rho(\mathbf{r})$, in the valence shell region of an atom (VSCC) are intimately related to the formation of chemical bonds and consequently to the presence of share electrons and/or lone pairs.^{119,121} This is particularly true in the case of metallic elements where the Laplacian distribution, in contrast to the charge density, can often reveal asphericity resulting from an incomplete filling of the *d*-shell, suggesting possible mechanisms for metal-ligand interactions.^{122,123}

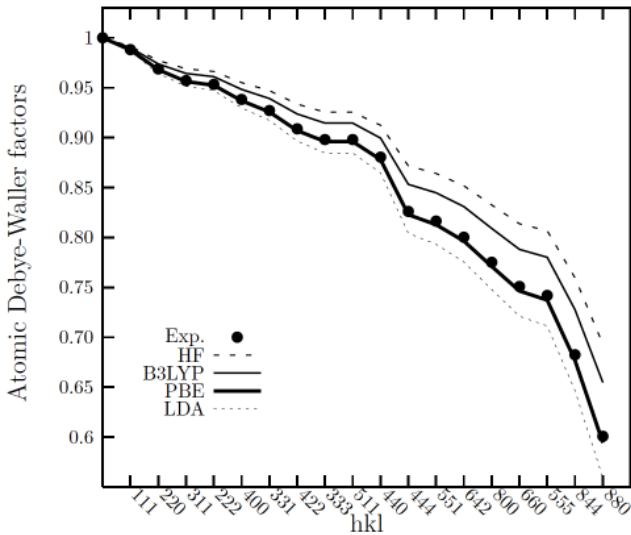


FIG. 17: Atomic Debye-Waller damping factors at 298 K, computed with four different Hamiltonians from their respective best determinations of the ADP for the set of 18 structure factors of Ref. 127 for silicon. The experimental points correspond to the ADP of Ref. 128.

Furthermore, a comparison of crystalline CPs and Laplacian features with the corresponding ones in the case of isolated molecules, or atoms, enables to evaluate packing and cooperative effects on the bonding character.¹²⁴ Moreover, QTAIMAC allows for the determination of atomic basins and their properties like the volume, the electronic population, Lagrangian and Hamiltonian electrons kinetic energy. Properties for the total system can be thus defined and calculated in terms of atomic contributions.¹²⁵ Finally, TOPOND can evaluate many local quantities like kinetic energy densities, virial densities and Becke electron localization function.¹²⁶ As a matter of fact, topological properties have been extensively used in the characterization of the energetic features of intermolecular interactions in molecular and weakly bound solids and have provided new descriptors and tools for the study of chemical bonding.

In Figure 16 the electron density, its Laplacian and the gradient trajectories are reported for two layered periodic hexagonal crystals, namely graphene (left panels) and boron nitride (right panels). Both structures can be understood as a network of strong interactions where the difference between homo- and hetero-bonds can be fully appreciated. The possibility given by CRYSTAL14 to incorporate electron correlation effects, by using hybrid, double-hybrid or MP2 corrected density matrices, represents a great opportunity and a significant progress towards a better understanding of the experimental results (see Section IX).

B. Debye-Waller Thermal Factors

Due to the fact that core and inner-valence electrons of atoms follow the movement of the respective nuclei, when ECD and related X-ray structure factors (XSF) are considered, it is mandatory to account for the effect of finite temperature, for instance by means of atomic harmonic Debye-Waller thermal factors.¹¹² An enormous amount of literature has been devoted to the approximate evaluation of such effects in order to allow for a correct interpretation of the X-ray scattering data.¹¹⁸

If a harmonic lattice potential is considered, then the probability density function of the nuclear displacements with respect to the equilibrium configuration of the atoms turns out to be a Gaussian function.¹²⁹ The most common way nuclear motion effects are dealt with when X-ray diffraction is considered is by means of Debye-Waller (DW) atomic factors which damp the diffraction intensities with respect to increasing wave number and temperature. Atomic DW factors are usually computed from atomic anisotropic displacement parameters (ADP). It has recently been suggested that ADPs are scarcely affected by anharmonicity so that harmonic mean-square displacements already provide a good description even of strongly anharmonic nuclear potentials.¹³⁰

In the CRYSTAL14 program, we have developed a fully *ab initio* approach for the computation of ADPs, by solving the lattice dynamics of the system (see Section VIC for details), DW factors and dynamical XSFs.²⁹ This scheme has been applied to the calculation of ADPs of a series of molecular crystals such as urea, benzene, urotropine and L-alanine and a satisfactory agreement has been reported with available experimental data.³⁰

As a test-case for the *ab initio* calculation of dynamical X-ray structure factors, we have considered crystalline silicon.²⁹ In general, experimental diffraction intensities and charge densities are less accurate than energy-related properties.¹³¹ Crystalline silicon represents an exception because of the high level of purity of its single crystals and availability of a very accurate technique for the measurement of dynamical structure factors (Pendellösung fringes method)^{132–134} which are known by an order of magnitude more accurately than for any other crystal.¹²⁷

In Figure 17 we report the DW damping factors computed with four different Hamiltonians from their respective best determinations of the ADPs for the set of 18 structure factors F_{hkl} of Ref. 127. The experimental points correspond to the ADP of Ref. 128. It is seen that HF and B3LYP underestimate the ADP and give a too small damping, the LDA overestimates the ADP while PBE is in very good agreement with experiment.

The technique used relies on the accurate description of the lattice dynamics and of the electron charge distribution of the system. The description of both aspects is dramatically affected by the adopted quantum chemical method. In the case of crystalline silicon, we find that the PBE functional of the DFT provides the best values

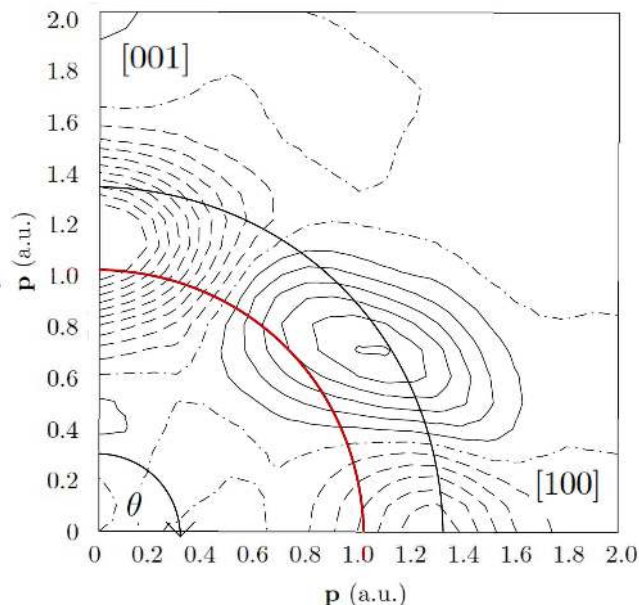


FIG. 18: (color online) Anisotropy map of the EMD of α -quartz as computed at HF level in a vertical plane; the vertical ($\theta=0^\circ$) and horizontal ($\theta=90^\circ$) axes coincide with the [001] and [100] crystallographic directions, respectively. The step between isolines is 0.01 a.u. and the maximum and minimum values reported are 0.1 and -0.1 a.u. The red thick and the black thin lines mark the region of the map where $|\mathbf{p}|=1.0$ and 1.3 a.u., respectively.

for both properties. An overall agreement factor of 0.47 % between the *ab initio* predicted values and the experimental determinations is found, as regards dynamical structure factors.

C. Momentum Density

The EMD $\pi(\mathbf{p})$ is a single-center function, invariant under the symmetry operations of the point group of the crystal, augmented with the *inversion* arising from the equality $\pi(\mathbf{p}) = \pi(-\mathbf{p})$;¹³⁵ such an object is a function of the counterintuitive momentum-space coordinates and it is characterized by a “collapsed” character about the origin $\mathbf{p} = \mathbf{0}$. For these reasons it is generally difficult to extract the information content of the EMD that is usually revealed in its very subtle features.

In recent years, a series of strategies have been devised and implemented in the CRYSTAL14 program for the analysis of the EMD and Compton profiles (CP) of crystalline materials:

- The computation of the EMD from the density matrix of the system; merits and drawbacks of this scheme with respect to that based on the crystalline orbital coefficients have been illustrated.³⁵ One of the main advantages is that of making the computation of the EMD of crystals possible at the MP2

level of theory, through the MP2 density matrix provided by the CRYSCOR program.¹³⁶

- The automated evaluation of the *spherical average* (SA) EMD function $\pi_{SA}(|\mathbf{p}|)$ and computation of the EMD-anisotropy $\Delta\pi(\mathbf{p}) = \pi(\mathbf{p}) - \pi_{SA}(|\mathbf{p}|)$. This scheme allows for the automated computation of EMD-anisotropy maps.³⁴ As an example, Figure 18 reports an anisotropy map of the EMD of α -quartz, as computed at HF level in the (010) plane. The region of maximum anisotropy lies at $|\mathbf{p}|$ values between 1.0 and 1.3 a.u. in this case.
- A partition scheme of the total EMD of a crystal into contributions coming from well-defined chemical subunits [$\pi(\mathbf{p}) = \sum_i \pi_i(\mathbf{p})$] that takes advantage of the localization of the crystalline orbitals into Wannier functions.^{34,137–141}
- As regards momentum space, the usual approximation of completely neglecting nuclear motions can result in a reasonable estimate of the CPs and the effect of finite temperature is seldom explicitly considered, although few *ad hoc* models have been reported.^{142,143} We have recently developed an *ab initio* Monte Carlo technique for the determination of the thermally averaged electronic first-order density matrix of crystals, in a harmonic approximation. The CPs and EMD of crystals can be computed at any temperature within such a scheme in a general, even if costly, way.¹⁴⁴
- The recent development of an algorithm for computing CPs from the density matrix of the system, rather than from the crystalline orbitals, made possible the investigation of the effect of the adopted computational method on momentum space properties, even beyond the one-electron approximation, with the MP2 approach implemented in the CRYSCOR program.¹³⁶ We have shown that very accurate directional CPs, as can be measured from the inelastic scattering of high intensity synchrotron radiation by single-crystal samples,^{145,146} can reveal subtle aspects of the electronic structure of periodic systems. In the cases of urea,¹⁴⁷ silicon^{148,149} and quartz³⁴ the theoretical CPs obtained using single-determinantal approximations to the ground-state wave function were found in fact to present definite discrepancies with respect to the experiment, which were partly removed when use was made of an *ab initio* technique based on a multi-determinantal description of the wave function, namely MP2.

IX. CLIMBING THE JACOB’S LADDER FOR SOLIDS

Density Functional Theory (DFT), in its Kohn-Sham formalism, has now reached a widespread success with

an ever increasing number of applications in chemistry, materials science and solid-state physics.¹⁵⁰ In the quest for the unknown exact exchange-correlation (XC) functional, many different approximations (DFA) have been proposed. Even if not systematically improvable, DFA can be classified in a hierarchical fashion according to the “Jacob’s Ladder” proposed by J. P. Perdew.¹⁵¹ When climbing the ladder, more complex ingredients are included in the mathematical form of the XC functional with the aim of reaching higher accuracy. The rungs of increasing accuracy/complexity are: (1) LDA, (2) GGA, (3) meta-GGA, (4) hyper-GGA and (5) RPA-like functionals.

In CRYSTAL14, all rungs of the Jacob’s Ladder are available. Along with LDA and GGA functionals, already included in the previous versions of the code, XC functionals belonging to the third-, fourth- and fifth-rung have been implemented, namely: semilocal and hybrid mGGA functionals, Range-Separated Hybrids (RSH) (4th-rung) and Double Hybrids (5th-rung).

A. New DFT functionals in CRYSTAL14

1. mGGA functionals

mGGA functionals currently implemented in CRYSTAL14 are τ -dependent and belong to the Minnesota’s

set, namely the M05 and M06 families.^{152–156} They include: the M05 global hybrid functional and its M05-2X variant with a doubled amount of HF exchange; the M06 hybrid functional and its variants from the 100% HF exchange (M06-HF) to the M06-2X functional with twice amount of exact exchange and the pure mGGA functionals M06-L.

2. Range-Separated Hybrid functionals

While global hybrids (GH) include a constant amount of exact exchange, in the RSH functionals the amount of HF exchange depends on the distance between electrons. This is obtained from the separation of the Coulomb operator in different ranges, usually by means of the *error function*.¹⁵⁷ When a partition into three pieces is adopted the Coulomb operator looks like:

$$\frac{1}{r_{12}} = \underbrace{\frac{\text{erfc}(\omega_{SR}r_{12})}{r_{12}}}_{SR} + \underbrace{\frac{1 - \text{erfc}(\omega_{SR}r_{12}) - \text{erf}(\omega_{LR}r_{12})}{r_{12}}}_{MR} + \underbrace{\frac{\text{erf}(\omega_{LR}r_{12})}{r_{12}}}_{LR} \quad (37)$$

where ω is the length scale of separation. The general form of a range-separated hybrid is then:

$$E_{xc}^{RSH} = E_{xc}^{DFA} + c_{SR}(E_{x,SR}^{HF} - E_{x,SR}^{DFA}) + c_{MR}(E_{x,MR}^{HF} - E_{x,MR}^{DFA}) + c_{LR}(E_{x,LR}^{HF} - E_{x,LR}^{DFA})$$

According to the values of c_{SR} , c_{MR} , c_{LR} , ω_{SR} and ω_{LR} , short-, middle- and long-range corrected RSH functionals can be defined. This allows one to include exact exchange in the selected interelectronic range and take advantage of its peculiar features. Short-range corrected (or Screened Coulomb) RSH functionals remove long-range HF exchange and are designed for solids where the LR-HF can lead to numerical instability, in particular for metallic systems. On the contrary, long-range corrected RSH functionals include HF exchange at LR to recover the correct decay of the exchange potential which is wrong in semilocal DFT functionals. In between, middle-range corrected RSH functionals are targeted to take advantage of the best of both worlds.

CRYSTAL14 offers a wide variety of RSH functionals, namely: HSE06¹⁵⁸ and HSEsol¹⁵⁹ (Short-range Corrected); HISS^{160,161} (Middle-range Corrected); RSHX LDA¹⁶², LC- ω PBE, LC- ω PBEsol¹⁶³, ω -B97 and ω B97-X^{164,165} (Long-range Corrected). The Henderson-Janesko-Scuseria model of the PBE exchange hole¹⁶⁶ has been adopted for the implementation of HSE06, HISS, LC- ω PBE and related RSH methods. This hole allows a fully analytical evaluation of the range-separated enhancement factor and recovers the correct PBE limit.

3. Double Hybrid functionals

Double hybrid functionals (DH) are hybridized not only in the exchange part by including a certain amount of HF exchange (i.e. the dependence on occupied orbitals) but also in the correlation part. In this case, hybridization involves the mixing with a contribution to the correlation energy that depends on unoccupied orbitals

through a MP2-like perturbative correction. Different schemes have been proposed, see Ref. 167 for a detailed list of double hybrid methods. Double hybrid functionals implemented in the code have the general formula:

$$E_{xc}^{DH} = (1-A)*E_x^{DFA} + A*E_x^{HF} + (1-B)*E_c^{DFA} + B*E_c^{MP2}$$

as proposed by Grimme.¹⁶⁸ A rigorous proof of the equation above within the adiabatic connection formalism has been recently reported.¹⁶⁹ The MP2-like correlation correction to the SCF energy is computed by means of the CRYSCOR program.^{136,170,171} According to the equation above, the following DH functionals have been made available in CRYSTAL14: B2PLYP, mPW2PLYP^{168,172} and B2GP-PLYP.¹⁷³ Although the cost of the calculation is the same as for MP2, double hybrids are less basis set dependent and partly overcome the drawbacks of DFT methods in the description of weak dispersive interactions.

4. Availability

Energies and analytic gradients are available for LDA, GGA, mGGA and related hybrid functionals, for both closed- and open-shell systems. DFT calculations can be run either sequentially or in parallel. For double-hybrids, only energies are available. Note that they have the same computational cost as MP2 rather than DFT. Calculations can only be run sequentially and for closed-shell systems. Parallel calculations will be made available with a future CRYSCOR release.¹⁷⁴

Approximately, the cost of the calculation is: LDA < GGA ~ mGGA < GH < RSH << DH. The higher cost of RSH with respect to GH is mainly due to the missing bipolar expansion approximation for two-electron exchange integrals.

Properties that require the solution of the CPKS equations (e.g. polarizability, hyper-polarizabilities and Raman intensities) are limited to a subset of LDA, GGA and related global hybrid functionals (see Ref. 36 for details).

B. Validation of XC functionals for solids

Most of the exchange-correlation functionals added to CRYSTAL14 have been designed and optimized for molecules. It is then interesting to validate them for application to solids. Moreover, it has been recently claimed that global hybrids and long-range corrected RSH functionals are impractical for calculation of crystalline systems.¹⁷⁵ Here, we show that they can be applied to solids as well. As a comparison between various DFA flavors, we report results for the prediction of the lattice parameter, bulk modulus and band gap of a representative set of nine simple solids, namely: three ionic solids which feature medium-to-wide band gap (i.e. LiF,

TABLE VII: Mean deviation (MD) and mean absolute deviation (MAD) for the equilibrium lattice constants (Å), bulk moduli (GPa) and band gaps (eV) from a variety of DFA with respect to reference values (see text for details) for the set of nine solids ^a

Method	Lattice constant		Bulk modulus		Band gap	
	MD	MAD	MD	MAD	MD	MAD
HF	0.062	0.065	16.7	17.6	6.32	6.32
LDA	-0.030	0.032	8.7	9.2	-1.84	1.84
PBE	0.082	0.082	-9.9	0.1	-2.01	2.01
PBEsol	0.028	0.029	-1.8	4.5	-1.98	1.98
B97	0.074	0.074	-4.6	5.6	-0.46	0.72
B3LYP	0.079	0.079	-5.7	7.2	-0.52	0.89
PBE0	0.032	0.035	6.1	6.7	-0.07	0.63
PBEsol0	-0.004	0.015	12.0	12.0	-0.02	0.62
HSE06	0.035	0.038	5.0	6.4	-0.69	0.71
HSEsol	-0.002	0.016	11.1	11.1	-0.65	0.65
HISS	0.007	0.022	16.3	16.3	0.13	0.51
LC- ω PBE	-0.007	0.036	21.0	21.0	4.28	4.28
LC- ω PBEsol	-0.053	0.053	35.0	35.0	5.49	5.49
ω -B97	0.027	0.028	11.2	11.2	4.22	4.22
ω -B97X	0.036	0.037	8.8	8.9	3.79	3.79
RSHXLDA	-0.009	0.037	17.9	18.2	4.50	4.50
M06-L	0.043	0.043	-0.9	6.8	-1.39	1.39
M06	0.051	0.056	1.2	6.3	-0.06	0.71
M06-2X	0.020	0.046	12.1	12.1	1.91	1.91

^a The full set of results is available as supporting information.

NaCl, MgO) and six semiconductors (i.e. C, Si, Ge, SiC, GaN, GaAs) ranging from very wide to very narrow band gap semiconductors.

Computed results are listed in Table VII. Mean deviations (with respect to a dataset of reference values are reported for eighteen XC functionals that belong to the first four rungs of the Jacob's Ladder. Results for HF are also included for comparison. Apart from LiF and NaCl, basis sets for other solids have been taken from Ref. 176. For LiF and NaCl a triple-zeta quality basis set has been employed. The reference dataset includes: (i) experimental lattice constants corrected for the zero-point anharmonic expansion, as reported in Ref. 177; experimental bulk moduli from Ref. 178 and (iii) low temperature (< 77 K) experimental (fundamental) band gaps.¹⁷⁹⁻¹⁸¹

As regards the prediction of the lattice parameters, PBEsol0 and HSEsol are the best performers, followed by HISS and PBEsol. As expected, HF tends to overestimate the lattice parameters, while LDA underestimates them. Pure GGA and mGGA functionals overestimate lattice constants but results improve when a GGA functional devised for solids, as PBEsol, is used. Inclusion of HF exchange in GHs and RSHs also improves the results but this depends on the adopted XC functional (e.g. PBE vs PBEsol based hybrids).

Most of the tested XC functionals tend to overestimate the bulk modulus of the examined solids. Oddly, the overestimation increases when passing from SC- to MC- to LC-RSH methods, as in the series: HSE06, HISS and LC- ω PBE. A possible explanation is that the reference data are less reliable than for lattice constants. In any case, PBEsol shows the lowest MAD, followed by B97, HSE06 and M06.

The correct prediction of the band gap is still a matter of debate in DFT. Here, we computed the band gap within the Kohn-Sham formalism as the difference between the top of the valence bands and the bottom of the conduction bands. This corresponds to the definition of the fundamental gap. Therefore, reference data have been chosen to be consistent to that definition. The examined systems span a range from 0.5 to 15 eV. As expected, all semilocal functionals underestimate the band gap, while global hybrids, short- and middle-range RSH functionals give very good results. On the contrary, LC-RSH methods tend to overestimate the band gap of solids. This is not unexpected because the inclusion of HF exchange at long-range makes them closer to HF which is known to largely overestimate the band gap. However, finding the best XC functional for band gap prediction is not an easy task, as recently discussed by some of us.¹⁸²

Overall, hybrid functionals, both global and short-range corrected, give the best results for basic solid state properties, in particular when combined with a semilocal functional devised for solids, as PBEsol. Nevertheless, a middle-range corrected functional as HISS might be a good compromise, because of its good performance for molecular properties as well.^{161,183} The good accuracy of hybrid HF/DFT methods as testified by Table VII, as well as other results reported in the present work, and their very efficient implementation in the code make then CRYSTAL14 stand out for its application in solid state chemistry and materials science.

X. CRYSTAL14 PERFORMANCE

In the preparation of CRYSTAL14 some effort has been devoted to improve general efficiency and optimize memory storage, particularly concerning executions in parallel on a large number of processors for large unit cell cases. Performance of CRYSTAL14 has already been discussed in the particular case of highly symmetric systems such as the nanotubes illustrated in Section II. Such calculations, which are shown to be feasible on one processor of an ordinary desktop, can be run very efficiently in parallel on a large number of processors because very good load-balancing in the computation of one- and two-electron integrals has now been complemented with an effective distribution of the matrix diagonalization task over the irreducible representations of the symmetry group of the system, that is, for example, 640 in the case of the largest nanotube considered in Section II. Indeed, one can

achieve almost ideal load-balancing and scaling in these particular cases where the nanotube symmetry properties imply that all irreducible representations include an equal number of basis functions, *i.e.* all matrices to be diagonalized have equal size.

Surfaces and interfacial phenomena, defective solids, biomaterials, and nanoparticulate systems, all require models including a large number of atoms in the unit cell and low symmetry, mostly P1. CRYSTAL14 has been optimized to be efficient and low memory consuming also in these cases and take advantage of the availability of High Performance Computing (HPC) resources, which have become the method of choice in those areas of science and technology that require the treatment of large amounts of data or the accomplishment of particularly demanding computational tasks. MPPCRYSTAL, the massive parallel version of the program, first available in CRYSTAL09, make it possible to perform calculations on hundreds or even thousands of CPUs, depending on system size. Scalability of the code both with respect to system size and with respect to the number of processors has been increased significantly with respect to previous implementation in CRYSTAL09.^{184,185} Efforts have been devoted also to the reduction of the memory footprint of the code, as the general trend in HPC systems is towards a decrease of the per core available RAM, due to the cost of memory both in terms of production and power consumption.

We refer the mesoporous silica MCM-41 structure as a test case for performance analysis of the current version of MPPCRYSTAL. It consists of a unit cell containing 579 atoms. The basis set was obtained from a 6-31G* Pople's standard basis set leading to nearly 8000 (7756) AOs in the unit cell. For tests at higher system sizes, supercells were built expanding the MCM-41 unit cell along the *c* axis. Different supercells will be denoted as X_j , with *j* being an integer defining the order of expansion along *c*. The B3LYP hybrid exchange-correlation density functional was used. Every run consists in an energy and gradient calculation, SCF+G, which are representative of a typical CRYSTAL run. For example, structure optimizations and phonon calculations consist essentially of several iterations of SCF+G steps.

A. CPU Time Scaling

MCM-41/ X_1 can be considered a medium-size system relative to current capabilities of MPPCRYSTAL since a SCF+G calculation can be run fairly efficiently on a limited number of cores (less than 10) and use of HPC resources is not strictly necessary, although fast-communication hardware is beneficial for attaining excellent scalability. SCF+G calculations for MCM-41/ X_1 were run in the range of 8-64 cores both on a local Linux cluster of Intel-Xeon processors with Ethernet connections and SuperMUC (LRZ, Germany), a HPC IBM System x iDataPlex powered by 16 Intel cores per node running at 2.7 GHz, with 2 GB/core (Figure 19).

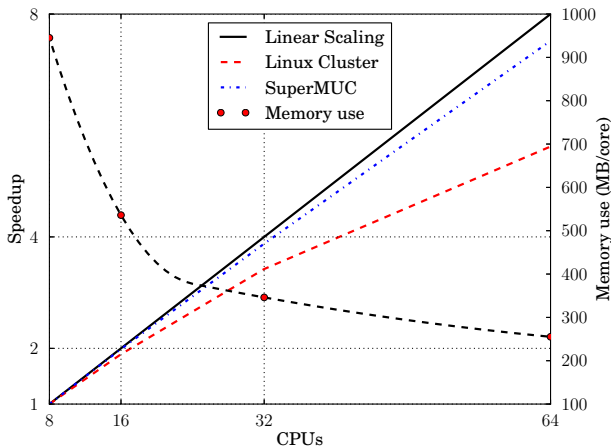


FIG. 19: Per core running time scaling and memory use with number of computing cores for MCM-41/X1 (7756 atomic orbitals in the unit cell) on a local Linux cluster and HPC SuperMUC system.

Scalability with the number of cores employed (n) is represented in terms of Speed-up with reference to the case with $n=8$: $\text{Speedup} = t_8/t_n$, where t_n is wall-clock time required for a SCF+G calculation run with n cores. The curve obtained with the HPC system closely approaches the bisectant in Figure 19 corresponding to ideal full scalability within this range of n . Out of this range, scalability is still good. For example, $\text{Speed-up} = 14.1$ *vs* an ideal value of 16 at $n=128$, but it tends to deteriorate for higher n (Speed-up 25.7 and 42.6 *vs* ideal 32 and 64, respectively, for $n=256$ and 512). The matrix diagonalization step carries the main responsibility for degrading scalability. In our experience the ratio of the number of basis functions in the unit cell (n_{BS}) to n is a good index for evaluating the expected degree of scalability and it can be stated that good scalability is granted by $n_{BS}/n \geq 50$. This criterion suggests that as many as $n=155$ cores can be used to run this particular calculation efficiently.

The other curve in Figure 19 obtained from performing the same calculations with the local cluster shows at what extent slow communications affect scalability. A Speed-up of 5.6 for $n=64$ is still not dramatically far from the ideal value of 8, though much worse than 7.5 as obtained with SuperMUC, thus suggesting that MP-PCRYSTAL can also be run on ordinary hardware without dramatic loss in performance. In absolute terms, it took 3970 wall-clock sec to run one full SCF+G CRYSTAL calculation for MCM-41/X1 on 64 cores of our local cluster *vs* 1990 sec on better performing SuperMUC. Recent improvements in memory storage management enabled us to run such a calculation also with PCRYSTAL by keeping memory requirements below 2 GB memory occupation per core. Comparison with PCRYSTAL emphasizes the great advantage of MPPCRYSTAL in this case. Indeed, this is a very unfavorable run with PCRYSTAL, because of expected large memory requirements and severe inefficiency in diagonalization.

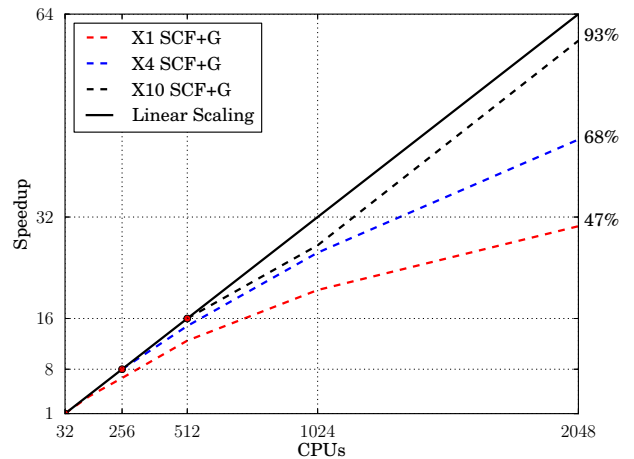


FIG. 20: Running time scaling with the number of computing cores as a function of the system size on the SuperMUC system. Different supercells of MCM-41 are considered, starting from X1 (7756 AOs) up to X10 (77560 AOs).

Diagonalization is performed only for matrices in the Γ point in this case, due to the large size of the unit cell. As a consequence, that step is thoroughly in charge of one of the cores with PCRYSTAL, all others being idle, and an SCF+G calculation required 74375 sec to be completed by 64 cores of our local cluster, that is, nearly 20 times more inefficiently than MPPCRYSTAL.

Data in Figure 19 and previous considerations suggest that MPPCRYSTAL can be used for even larger unit cell cases provided an adequate amount of computational resources is allocated. That such a possibility exists is confirmed by data reported in Figure 20, where performance in SCF+G calculations for supercells of MCM-41 up to MCM-41/X10 is shown in a range of 32-2048 core utilization (in this case Speedup is referred to t_{32}). In the first place, Figure 20 verifies that very large unit-cell calculations can be run in the scale of nowadays easily accessible HPC resources. Moreover, CRYSTAL14 appears to be fairly efficient and shows very good scalability for any appropriate choice of n with respect to the size of the problem. That is easily appreciated, for example, by comparing curves referring to MCM-41/X1 with MCM-41/X10, as best scalability, satisfying the criterion defined above based on the ratio of n_{BF} to n , is obtained for $n \leq 155$ and $n \leq 1550$, respectively, in the two cases.

Very similar trends were obtained by running the same calculations on other HPC architectures, such as HEC-ToR (UK), a Cray XE6 system, based on 12-core AMD Opteron 2.1GHz Magny Cours processors, and Fermi (CINECA, Italy), an IBM Blue Gene Q system. Scalability improves slightly for IBM Blue Gene systems, because a poorer per core performance implies a smaller visibility of the communication overheads and a less relevance of the worst scaling algorithms such as the diagonalization step. It is important to notice that scalability can depend strongly on various factors, since different parts of

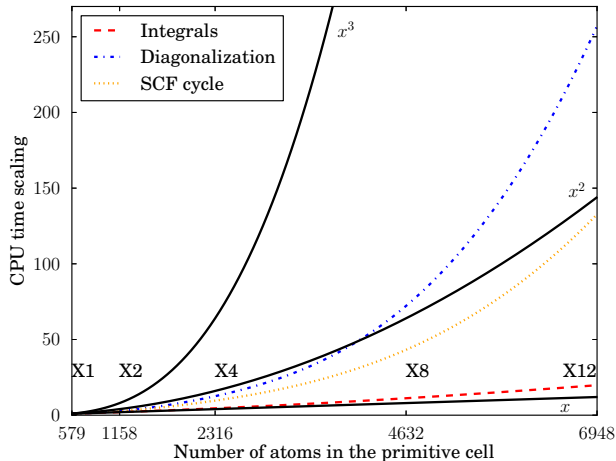


FIG. 21: Running time scaling for a SCF calculation with increasing size of the supercell of MCM-41 on 1024 cores of the HPC SuperMUC system. Scaling of single steps such as the calculation of the integrals and diagonalization are also reported with reference to ideally linear, quadratic and cubic trends.

the code show better parallel performance with respect to others. For example, if on one hand the two-electron integral algorithm is known to scale perfectly with the number of CPUs, on the other hand the diagonalization procedure, which is based on the Scalapack routines, shows a rather poor scaling. The overall scaling depends on the relative importance of these two steps in the SCF cycle. This is apparent in Figure 21 where, despite the good scaling previously shown with the number of cores, computational time increases almost quadratically with the size of the supercell. Indeed, because the calculation of two-electron integrals has become still more efficient and closely approaches linear scaling in CRYSTAL14, time required for matrix diagonalization, which exhibits a quadratic behavior as far as MCM41/X8 (then tending to cubic), dominates and determines the general trend.

Furthermore, scaling depends strongly on the density of the system under investigation and the number of one- and two-electron integrals to be computed depends on the overlap of the basis functions. For the reasons mentioned above more compact crystals than MCM-41 are expected to exhibit better overall scaling with respect to the number of cores used. As an example, data about performance for a $3 \times 3 \times 3$ supercell of calcite is reported in table VIII. Although such a supercell contains less than half of the atoms in MCM-41/X1 (270 *vs.* 579), scalability on SuperMUC follows a better scaling: at $n=2048$, Speed-up is 43.5 for calcite and only 26.1 in the case of MCM-41/X1, both to be compared with an ideal factor of 64.

TABLE VIII: Speed-up for a $3 \times 3 \times 3$ calcite supercell as a function of the number of cores (n). All calculations were performed with the hybrid B3LYP functional on SuperMUC. t_n is wall-clock time in seconds. Efficiency was evaluated as: Speed-up $\times 32 \times 100/n$

n	t_n	Speedup	Efficiency
32	38471.3	1.00	100%
64	19769.0	1.95	97.3%
128	9866.3	3.90	97.4%
256	5116.4	7.52	93.9%
512	2753.6	13.97	87.3%
1024	1576.7	24.40	76.2%
2048	883.4	43.55	68.0%

B. Memory Storage

Diminishing memory requirements has been a main challenging issue in extending the capabilities of CRYSTAL to handle large unit-cell systems in the HPC context where the general trend implies reduced memory availability both in terms of GB/core and bandwidth. Such result has been achieved partially by extending the use of dynamical memory allocation enormously, optimizing memory usage throughout the entire code, particularly in the most memory consuming steps, carefully taking sparsity into account and progressively distributing more and more data to the cores.

Figure 19 contains a plot of the maximum memory required by MPPCRYSTAL to run a MCM-41/X1 on a different number of cores. One of the main features of MPPCRYSTAL is an optimized use of memory resources obtained by distributing data to cores as far as possible. However, even MPPCRYSTAL includes a part of replicated data, which are mainly related to information of very general use, such as mapping tables, or data that, if distributed according to a predetermined pattern, would downgrade load balancing of the algorithms. This is the reason why the memory-occupation curve in the Figure appears as to be tending to an asymptotic value for large values of n , which represents the maximum amount of replicated data being stored to memory during an SCF+G calculation for MCM-41/X1. The maximum request for memory amounts to 945 MB per core at $n=8$. Such request is drastically reduced at $n=16$ because memory occupation is dominated by the size of the matrices defined in the reciprocal space (Fock and overlap matrices, eigenvectors) which are fully distributed to the cores, so that a higher number of cores corresponds to more effectively distributed data. As n grows, blocks of such matrices distributed to each core become smaller and smaller until data distribution is almost complete and the base of the replicated-data emerges.

Figure 22 shows how memory request evolves when considering large supercells of MCM-41 and how effi-

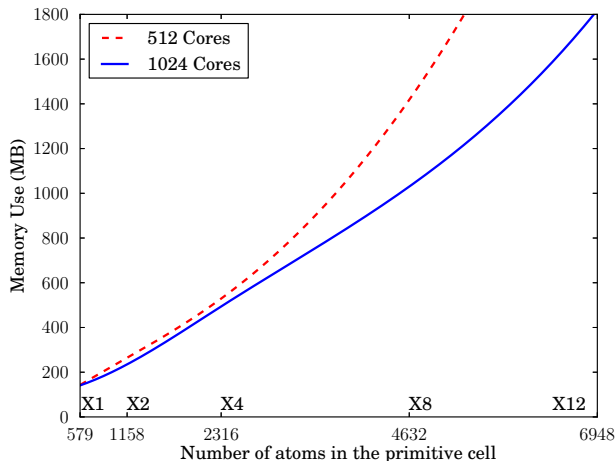


FIG. 22: Maximum per core memory usage (Megabytes) in a SCF calculation for MCM-41 supercells with increasing size on 512 and 1024 cores of the HPC SuperMUC system.

ciently increasing n enables to extend the size of the system that can be handled below an assumed threshold of 2 GB of memory used per core, which is about the amount of memory one expects to be available in most HPC systems, like SuperMUC. Figure 22 shows that, for example, supercells as large as MCM-41/12X satisfy such requirement with $n=2048$. For larger supercells, such as MCM-41/16X (more than 110000 basis functions), preliminary calculations show that the replicated-data part becomes relevant and approaches the 2 GB limit, so that

improvements are needed to further reduce memory requirements and enhance data distribution.

Acknowledgments

This paper is dedicated to the memory of Cesare Pisani (1938-2011) and Carla Roetti (1943-2010) who started the CRYSTAL project in the seventies and gave a fundamental contribution to the development of the code during all their scientific careers. Victor R. Saunders is acknowledged for his invaluable contribution in the development of the code and in making it computationally efficient and numerically stable.

Contribution to validate the new features by applying them to research problem is recognized to all people who have been working in the Theoretical Chemistry Group of the University of Torino since 2009. In particular, help to test the current version of the code is acknowledged to: Simone Salustro, Jacopo Baima, Marco Lorenz, Agnes Mahmoud, Valentina Lacivita, Gustavo Sophia, Elisa Albanese, Davide Presti, Massimo delle Piane, Marta Corno, Raffaella Demichelis, Anna M. Ferrari and Giuseppe Mallia.

Authors would like to thank Piero Ugliengo for continuous help, useful suggestions, rigorous testing.

Improvements of the CRYSTAL code in its massive parallel version were made possible thanks to the PRACE proposals no. 2011050810 and 2013081680.

* Electronic address: crystal@unito.it

- ¹ C. Pisani, R. Dovesi, and C. Roetti, *Hartree-Fock Ab Initio Treatment of Crystalline Solids*, vol. 48 of *Lecture Notes in Chemistry Series* (Springer Verlag, Berlin, 1988).
- ² R. Dovesi, C. Pisani, C. Roetti, M. Causà, and V. Saunders, *CRYSTAL88*, QCPE Program No. 577, Indiana University, Bloomington, IN (1988).
- ³ R. Dovesi, *Int. J. Quantum Chem.* **29**, 1755 (1986).
- ⁴ C. Zicovich-Wilson and R. Dovesi, *Int. J. Quantum Chem.* **67**, 299 (1998).
- ⁵ C. Zicovich-Wilson and R. Dovesi, *Int. J. Quantum Chem.* **67**, 311 (1998).
- ⁶ C. M. Zicovich-Wilson, Y. Noël, A. M. Ferrari, R. Orlando, M. De La Pierre, and R. Dovesi, *AIP Conf. Proc.* **1456**, 248 (2012).
- ⁷ R. Orlando, M. De La Pierre, A. Erba, C. M. Zicovich-Wilson, and R. Dovesi (2014), in preparation.
- ⁸ W. F. Perger, J. Criswell, B. Civalieri, and R. Dovesi, *Comput. Phys. Commun.* **180**, 1753 (2009).
- ⁹ A. Erba, A. Mahmoud, R. Orlando, and R. Dovesi, *Phys. Chem. Minerals* **41**, 151 (2014).
- ¹⁰ A. Erba, A. Mahmoud, R. Orlando, and R. Dovesi, *Phys. Chem. Minerals* **41**, 161 (2014).
- ¹¹ A. Erba, M. Ferrabone, J. Baima, R. Orlando, M. Rérat, and R. Dovesi, *J. Chem. Phys.* **138**, 054906 (2013).
- ¹² A. Erba, K. E. El-Kelany, M. Ferrero, I. Baraille, and M. Rérat, *Phys. Rev. B* **88**, 035102 (2013).
- ¹³ A. Erba and R. Dovesi, *Phys. Rev. B* **88**, 045121 (2013).
- ¹⁴ M. Ferrero, M. Rérat, R. Orlando, and R. Dovesi, *J. Comput. Chem.* **29**, 1450 (2008).
- ¹⁵ M. Ferrero, M. Rérat, R. Orlando, and R. Dovesi, *J. Chem. Phys.* **128**, 014110 (2008).
- ¹⁶ M. Ferrero, M. Rérat, R. Orlando, R. Dovesi, and I. Bush, *J. Phys. Conf. Ser.* **117**, 12016 (2008).
- ¹⁷ M. Ferrero, M. Rérat, B. Kirtman, and R. Dovesi, *J. Chem. Phys.* **129**, 244110 (2008).
- ¹⁸ R. Orlando, M. Ferrero, M. Rérat, B. Kirtman, and R. Dovesi, *J. Chem. Phys.* **131**, 184105 (2009).
- ¹⁹ R. Orlando, V. Lacivita, R. Bast, and K. Ruud, *J. Chem. Phys.* **132**, 244106 (2010).
- ²⁰ F. Pascale, C. M. Zicovich-Wilson, F. L. Gejo, B. Civalieri, R. Orlando, and R. Dovesi, *J. Comp. Chem.* **25**, 888 (2004).
- ²¹ C. M. Zicovich-Wilson, F. Pascale, C. Roetti, V. R. Saunders, R. Orlando, and R. Dovesi, *J. Comp. Chem.* **25**, 1873 (2004).
- ²² L. Maschio, B. Kirtman, M. Rérat, R. Orlando, and R. Dovesi, *J. Chem. Phys.* **139**, 164101 (2013).
- ²³ L. Maschio, B. Kirtman, M. Rérat, R. Orlando, and R. Dovesi, *J. Chem. Phys.* **139**, 164102 (2013).
- ²⁴ L. Maschio, B. Kirtman, R. Orlando, and M. Rérat, *J. Chem. Phys.* **137**, 204113 (2012).

- ²⁵ L. Maschio, B. Kirtman, M. Rérat, R. Orlando, and R. Dovesi, *J. Chem. Phys.* **139**, 167101 (2013).
- ²⁶ M. De La Pierre, R. Orlando, L. Maschio, K. Doll, P. Ugliengo, and R. Dovesi, *J. Comp. Chem.* **32**, 1775 (2011).
- ²⁷ A. M. Ferrari, L. Valenzano, A. Meyer, R. Orlando, and R. Dovesi, *Journal of Physical Chemistry A* **113**, 11289 (2009).
- ²⁸ C. Carteret, M. De La Pierre, M. Dossot, F. Pascale, A. Erba, and R. Dovesi, *J. Chem. Phys.* **138**, 014201 (2013).
- ²⁹ A. Erba, M. Ferrabone, R. Orlando, and R. Dovesi, *J. Comput. Chem.* **34**, 346 (2013).
- ³⁰ A. O. Madsen, B. Civalleri, M. Ferrabone, F. Pascale, and A. Erba, *Acta Crystallogr. Sec. A* **69**, 309 (2013).
- ³¹ S. Mustapha, P. D'Arco, M. De La Pierre, Y. Noel, M. Ferrabone, and R. Dovesi, *J. Phys.: Condens. Matter* **25**, 105401 (2013).
- ³² P. D'Arco, S. Mustapha, M. Ferrabone, Y. Noel, M. De La Pierre, and R. Dovesi, *J. Phys.: Condens. Matter* **25**, 355401 (2013).
- ³³ C. Gatti, *TOPOND-96 : an electron density topological program for systems periodic in N (N=0-3) dimensions, User's manual*, CNR-CSR SRC, Milano (1996).
- ³⁴ C. Pisani, A. Erba, S. Casassa, M. Itou, and Y. Sakurai, *Phys. Rev. B* **84**, 245102 (2011).
- ³⁵ A. Erba and C. Pisani, *J. Comput. Chem.* **33**, 822 (2012).
- ³⁶ R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, K. Doll, N. M. Harrison, B. Civalleri, I. J. Bush, et al., *CRYSTAL14 User's Manual*, Università di Torino, Torino (2014), <http://www.crystal.unito.it>.
- ³⁷ Tutorials on how to run specific types of calculations with CRYSTAL can be found at <http://www.crystal.unito.it/tutorials>.
- ³⁸ Y. Noël, P. D'Arco, R. Demichelis, C. M. Zicovich-Wilson, and R. Dovesi, *J. Comput. Chem.* **31**, 855 (2010).
- ³⁹ Y. Noël, M. De La Pierre, C. Z. Wilson, R. Orlando, and R. Dovesi (2014), in preparation.
- ⁴⁰ R. Credendino, V. Busico, M. Causà, V. Barone, P. H. M. Budzelaar, and C. Zicovich-Wilson, *Phys. Chem. Chem. Phys.* **11**, 6525 (2009).
- ⁴¹ M. Branda, R. Ferullo, M. Causà, and F. Illas, *J. Phys. Chem. C* **115**, 3716 (2011).
- ⁴² W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The art of scientific computing* (University Press, New York, 1989).
- ⁴³ J. Simons, P. Jørgensen, H. Taylor, and J. Ozment, *J. Phys. Chem.* **87**, 2745 (1983).
- ⁴⁴ P. Pulay, G. Fogarasi, F. Peng, and J. E. Boggs, *J. Am. Chem. Soc.* **101**, 2550 (1979).
- ⁴⁵ P. Pulay, *Chem. Phys. Lett.* **100**, 151 (1983).
- ⁴⁶ C. Peng, P. Y. Ayala, H. B. Schlegel, and M. J. Frish, *J. Comput. Chem.* **17**, 49 (1996).
- ⁴⁷ P. Pulay and G. Fogarasi, *J. Chem. Phys.* **96**, 2856 (1992).
- ⁴⁸ K. N. Kudin, G. E. Scuseria, and H. B. Schlegel, *J. Chem. Phys.* **114**, 2919 (2001).
- ⁴⁹ A. Rimola, C. M. Zicovich-Wilson, R. Dovesi, and P. Ugliengo, *J. Chem. Theor. Comput.* **6**, 1341 (2010).
- ⁵⁰ E. Polak, *Computational Methods in Optimizations* (Academic Press, New York, 1971).
- ⁵¹ Y. Noël, C. M. Zicovich-Wilson, B. Civalleri, P. D'Arco, and R. Dovesi, *Phys. Rev. B* **65**, 014111 (2001).
- ⁵² J. Baima, A. Erba, R. Orlando, M. Rérat, and R. Dovesi, *J. Phys. Chem. C* **117**, 12864 (2013).
- ⁵³ V. Lacivita, A. Erba, Y. Noël, R. Orlando, P. D'Arco, and R. Dovesi, *J. Chem. Phys.* **138**, 214706 (2013).
- ⁵⁴ K. E. El-Kelany, M. Ferrabone, M. Rérat, P. Carbonniere, C. M. Zicovich-Wilson, and R. Dovesi, *Phys. Chem. Chem. Phys.* **15**, 13296 (2013).
- ⁵⁵ M. U. Kahaly and U. V. Waghmare, *J. Nanosci. Nanotechnol.* **1787**, 2007 (7).
- ⁵⁶ E. Hernández, C. Goze, P. Bernier, and A. Rubio, *Phys. Rev. Lett.* **80**, 4502 (1998).
- ⁵⁷ M. J. P. Musgrave, *Crystal Acoustics* (Holden-Day, San Francisco, California, 1970).
- ⁵⁸ B. A. Auld, *Acoustic Fields and Waves in Solids* (Krieger Publishing Company, Malabar, Florida, 1973).
- ⁵⁹ B. B. Karki, L. Stixrude, and R. M. Wentzcovitch, *Rev. Geophys.* **39**, 507 (2001).
- ⁶⁰ F. Jiang, S. Speziale, S. R. Shieh, and T. S. Duffy, *J. Phys.:Cond. Matter* **16**, S1041 (2004).
- ⁶¹ D. L. Anderson, *Theory of the Earth* (Blackwell Scientific Publications, Boston, 1989).
- ⁶² S. M. Nakhmanson, A. Calzolari, V. Meunier, J. Bernholc, and M. B. Nardelli, *Phys. Rev. B* **67**, 235406 (2003).
- ⁶³ P. Panda, *J. Mater. Sci.* **44**, 5049 (2009), ISSN 0022-2461.
- ⁶⁴ S. Zhang, R. Xia, L. Lebrun, D. Anderson, and T. R. Shrout, *Mater. Lett.* **59**, 3471 (2005).
- ⁶⁵ D. E. Grupp and A. M. Goldman, *Science* **276**, 392 (1997).
- ⁶⁶ A. Mahmoud, A. Erba, K. E. El-Kelany, M. Rérat, and R. Orlando, *Phys. Rev. B* **89**, 045103 (2014).
- ⁶⁷ W. Cady, *Piezoelectricity*, vol. 1 (Dover Publications, 1964).
- ⁶⁸ J. Haines, O. Cambon, D. A. Keen, M. G. Tucker, and M. T. Dove, *Appl. Phys. Lett.* **81**, 2968 (2002).
- ⁶⁹ E. Philippot, P. Armand, P. Yot, O. Cambon, A. Goiffon, G. J. McIntyre, and P. Bordet, *J. Solid State Chem.* **146**, 114 (1999).
- ⁷⁰ V. Ranieri, S. Darracq, M. Cambon, J. Haines, O. Cambon, A. Largeau, and G. Demazeau, *Inorg. Chem.* **50**, 4632 (2011).
- ⁷¹ K. E. El-Kelany, A. Erba, P. Carbonniere, and M. Rérat, *J. Phys.: Cond. Matter* (2013), submitted.
- ⁷² A. A. Giardini and E. Poindexter, *J. Opt. Soc. Amer.* **48**, 556 (1958).
- ⁷³ K. V. Krishna Rao and V. G. Krishna Murty, *Acta Crystallogr.* **17**, 788 (1964).
- ⁷⁴ M. Cardona, W. Paul, and H. Brooks, *J. Phys. Chem. Solids* **8**, 204 (1959).
- ⁷⁵ M. Ferrero, B. Civalleri, M. Rérat, R. Orlando, and R. Dovesi, *J. Chem. Phys.* **131**, 214704 (2009).
- ⁷⁶ V. Lacivita, M. Rérat, B. Kirtman, M. Ferrero, R. Orlando, and R. Dovesi, *J. Chem. Phys.* **131**, 204509 (2009).
- ⁷⁷ V. Lacivita, M. Rérat, R. Orlando, M. Ferrero, and R. Dovesi, *J. Chem. Phys.* **136**, 114101 (2012).
- ⁷⁸ M. Ferrabone, B. Kirtman, R. Orlando, M. Rérat, and R. Dovesi, *Phys. Rev. B* **83**, 235421 (2011).
- ⁷⁹ A. Ferrari, M. Lessio, and M. Rérat, *Comp. Mat. Science* **68**, 280 (2013).
- ⁸⁰ C. S. Yannoni and T. C. Clarke, *Phys. Rev. Lett.* **51**, 1191 (1983).
- ⁸¹ H. Shirakawa, T. Ito, and S. Ikeda, *Polym. J.* **4**, 460 (1973).
- ⁸² H. Sekino, Y. Maeda, M. Kamiya, and K. Hirao, *J. Chem. Phys.* **126**, 014107 (2007).
- ⁸³ W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).

- ⁸⁴ S. H. Vosko, L. Wilk, and M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
- ⁸⁵ C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
- ⁸⁶ M. Causà, R. Dovesi, R. Orlando, C. Pisani, and V. R. Saunders, *J. Phys. Chem.* **92**, 909 (1988).
- ⁸⁷ R. Orlando, R. Bast, K. Ruud, U. Ekström, M. Ferrabone, B. Kirtman, and R. Dovesi, *J. Phys. Chem. A* **115**, 12631 (2011).
- ⁸⁸ G. Burns and F. Dacol, *Solid State Communications* **42**, 9 (1982).
- ⁸⁹ D. M. Bishop, M. Hasan, and B. Kirtman, *J. Chem. Phys.* **103**, 4157 (1995).
- ⁹⁰ M. Ferrabone, B. Kirtman, V. Lacivita, M. Rérat, R. Orlando, and R. Dovesi, *Int. J. Quantum Chem.* **112**, 2160 (2012).
- ⁹¹ K. Doll, N. M. Harrison, and V. R. Saunders, *Int. J. Quantum Chem.* **82**, 1 (2001).
- ⁹² K. Doll, *Comput. Phys. Comm.* **137**, 74 (2001).
- ⁹³ L. Maschio, B. Kirtman, S. Salustro, C. M. Zicovich-Wilson, R. Orlando, and R. Dovesi, *J. Phys. Chem. A* **117**, 11464 (2013).
- ⁹⁴ M. Prencipe, L. Maschio, B. Kirtman, S. Salustro, A. Erba, and R. Dovesi, *J. Phys. Chem. A* (2014), submitted.
- ⁹⁵ L. Maschio and R. Dovesi, *La chimica e l'industria* (2014), in press.
- ⁹⁶ H. Suto, H. Sogawa, S. Tachibana, C. Koike, H. Karoji, A. Tsuchiyama, H. Chihara, K. Mizutani, J. Akedo, K. Ogiso, et al., *Mon. Not. R. Astron. Soc.* **370**, 1599 (2006).
- ⁹⁷ G. M. Barrow, *Introduction to molecular spectroscopy* (McGraw-Hill, New York, 1962).
- ⁹⁸ B. A. Hess, L. J. Schaad, P. Carsky, and R. Zahradnik, *Chem. Rev.* **86**, 709 (1986).
- ⁹⁹ J. Decius and R. Hexter, *Molecular vibrations in crystals* (McGraw-Hill, 1977).
- ¹⁰⁰ M. Veithen, X. Gonze, and P. Ghosez, *Phys. Rev. B* **71**, 125107 (2005).
- ¹⁰¹ S. A. Prosandeev, U. Waghmare, I. Levin, and J. Maslar, *Phys. Rev. B* **71**, 214307 (2005).
- ¹⁰² W. D. Johnston, *Phys. Rev. B* **1**, 3494 (1970).
- ¹⁰³ J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, *J. Phys.-Condens. Mat.* **14**, 2745 (2002).
- ¹⁰⁴ G. Kresse and J. Furthmüller, *Comput. Mat. Sci.* **6**, 15 (1996).
- ¹⁰⁵ G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
- ¹⁰⁶ K. Parlinski, Z. Q. Li, and Y. Kawazoe, *Phys. Rev. Lett.* **78**, 4063 (1997).
- ¹⁰⁷ A. Togo, F. Oba, and I. Tanaka, *Phys. Rev. B* **78**, 134106 (2008).
- ¹⁰⁸ M. Ferrabone, A. Mahmoud, and R. Dovesi (2014), in preparation.
- ¹⁰⁹ H. T. Haselton and E. F. Westrum, *Geochimica et Cosmochimica Acta* pp. 701–709 (1980).
- ¹¹⁰ C. Tequi, R. A. Robie, B. S. Hemingway, D. Neuville, and P. Richet, *Geochimica et Cosmochimica Acta* **55**, 10051010 (1991).
- ¹¹¹ A. M. Hofmeister and A. Chopelas, *Am. Mineral.* **76**, 880 (1991).
- ¹¹² K. N. Trueblood, H.-B. Bürgi, H. Burzlaff, J. D. Dunitz, C. M. Gramaccioli, H. H. Shulz, U. Shmueli, and C. Abraham, *Acta Cryst. A* **52**, 770 (1996).
- ¹¹³ S. Swaminathan, B. M. Craven, and R. K. McMullan, *Acta Crystallogr. Sec. B* **40**, 300 (1984).
- ¹¹⁴ C. Pisani, R. Dovesi, A. Erba, and P. Giannozzi, in *Modern Charge Density Analysis*, edited by C. Gatti and P. Macchi (Springer, Berlin, 2011).
- ¹¹⁵ Y. Tanaka, Y. Sakurai, A. T. Stewart, N. Shiotani, P. E. Mijnen, S. Kaprzyk, and A. Bansil, *Phys. Rev. B* **63**, 045120 (2001).
- ¹¹⁶ P. Coppens and A. Volkov, *Acta Cryst.* **A60**, 357 (2004).
- ¹¹⁷ B. Schiott, J. Overgaard, F. K. Larsen, and B. B. Iversen, *Int. J. Quant. Chem.* **96**, 23 (2004).
- ¹¹⁸ T. S. Koritsanszky and P. Coppens, *Chem. Rev.* **101**, 1583 (2001).
- ¹¹⁹ R. F. W. Bader, *Atoms in Molecules - A Quantum Theory* (Oxford University Press, Oxford, UK, 1990).
- ¹²⁰ C. Gatti, *Zeitschrift für Kristallographie* **220**, 399 (2005).
- ¹²¹ P. Popelier, *Atoms in Molecules. An Introduction*. (Prentice Hall, Pearson Education Limited, 2000).
- ¹²² P. Macchi and A. Sironi, *Coordination Chem. Phys.* **238-239**, 383 (2003).
- ¹²³ R. Bianchi, C. Gatti, V. Adovasio, and M. Nardelli, *Acta Cryst.* **B52**, 471 (1996).
- ¹²⁴ C. Gatti, V. R. Saunders, and C. Roetti, *J. Chem. Phys.* **101**, 10686 (1994).
- ¹²⁵ C. Gatti, *Phys. Scr.* **87**, 048102 (2013).
- ¹²⁶ A. Becke and K. Edgecombe, *J. Chem. Phys.* **92**, 5397 (1990).
- ¹²⁷ Z. W. Lu, A. Zunger, and M. Deutsch, *Phys. Rev. B* **47**, 9385 (1993).
- ¹²⁸ C. Flensburg and R. F. Stewart, *Phys. Rev. B* **60**, 284 (1999).
- ¹²⁹ J. D. Dunitz, V. Schomaker, and K. N. Trueblood, *J. Phys. Chem.* **92**, 856 (1988).
- ¹³⁰ D. J. Safarik, A. Llobet, and J. C. Lashley, *Phys. Rev. B* **85**, 174105 (2012).
- ¹³¹ J. M. Zuo, P. Blaha, and K. Schwarz, *J. Phys.: Condens. Matter* **9**, 7541 (1997).
- ¹³² T. Saka and N. Kato, *Acta Cryst. A* **42**, 469 (1986).
- ¹³³ R. Teworte and U. Bonse, *Phys. Rev. B* **29**, 2102 (1984).
- ¹³⁴ P. J. E. Aldred and M. Hart, *Proc. R. Soc. London* **332**, 223 (1973).
- ¹³⁵ A. J. Thakkar, *Adv. Chem. Phys.* **128**, 303 (2004).
- ¹³⁶ C. Pisani, M. Schütz, S. Casassa, D. Usvyat, L. Maschio, M. Lorenz, and A. Erba, *Phys. Chem. Chem. Phys.* **14**, 7615 (2012).
- ¹³⁷ G. H. Wannier, *Phys. Rev.* **52**, 191 (1937).
- ¹³⁸ C. M. Zicovich-Wilson, R. Dovesi, and V. R. Saunders, *J. Chem. Phys.* **115**, 9708 (2001).
- ¹³⁹ S. Casassa, C. M. Zicovich-Wilson, and C. Pisani, *Theor. Chem. Acc.* **116**, 726 (2006).
- ¹⁴⁰ C. M. Zicovich-Wilson and A. Erba, *Theor. Chem. Acc.* **126**, 165 (2010).
- ¹⁴¹ C. M. Zicovich-Wilson and A. Erba, *Int. J. Quantum Chem.* **112**, 3543 (2012).
- ¹⁴² S. B. Dugdale and T. Jarlborg, *Solid State Commun.* **105**, 283 (1998).
- ¹⁴³ C. Sternemann, G. Döring, C. Wittkop, W. Schülke, A. Shukla, T. Buslaps, and P. Suortti, *J. Phys. Chem. Solids* **61**, 379 (2000).
- ¹⁴⁴ C. Pisani, A. Erba, M. Ferrabone, and R. Dovesi, *J. Chem. Phys.* **137**, 044114 (2012).
- ¹⁴⁵ Y. Sakurai, Y. Tanaka, A. Bansil, S. Kapzyk, A. T. Stewart, Y. Nagashima, T. Hyodo, S. Nanao, H. Kawata, and

- N. Shiotani, Phys. Rev. Lett. **74**, 2252 (1995).
- ¹⁴⁶ C. Blaas, J. Redinger, S. Manninen, V. Honkimäki, K. Hämäläinen, and P. Suortti, Phys. Rev. Lett. **75**, 1984 (1995).
- ¹⁴⁷ A. Erba, C. Pisani, S. Casassa, L. Maschio, M. Schütz, and D. Usvyat, Phys. Rev. B **81**, 165108 (2010).
- ¹⁴⁸ C. Pisani, M. Itou, Y. Sakurai, R. Yamaki, M. Ito, A. Erba, and L. Maschio, Phys. Chem. Chem. Phys. **13**, 933 (2011).
- ¹⁴⁹ A. Erba, M. Itou, Y. Sakurai, R. Yamaki, M. Ito, S. Casassa, L. Maschio, A. Terentjevs, and C. Pisani, Phys. Rev. B **83**, 125208 (2011).
- ¹⁵⁰ K. Burke, J. Chem. Phys. **136**, 150901 (2012).
- ¹⁵¹ J. P. Perdew and K. Schmidt, AIP Conf. Proc. **577**, 1 (2001).
- ¹⁵² Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Phys. **123**, 161103 (2005).
- ¹⁵³ Y. Zhao, N. E. Schultz, and D. G. Truhlar, J. Chem. Theory Comput. **2**, 364 (2005).
- ¹⁵⁴ Y. Zhao and D. G. Truhlar, J. Chem. Phys. **125**, 194101 (2006).
- ¹⁵⁵ Y. Zhao and D. G. Truhlar, Theor. Chem. Acc. **120**, 215 (2008).
- ¹⁵⁶ Y. Zhao and D. G. Truhlar, J. Phys. Chem. A **110**, 13126 (2006).
- ¹⁵⁷ A. Savin, *Recent development and applications of Density Functional Theory* (Elsevier, Amsterdam, 1996), p. 327.
- ¹⁵⁸ A. V. Krukau, O. A. Vydrov, A. F. Izmaylov and G. E. Scuseria, J. Chem. Phys. **125**, 224106 (2006).
- ¹⁵⁹ L. Schimka, J. Harl and G. Kresse, J. Chem. Phys. **134**, 024116 (2011).
- ¹⁶⁰ T. M. Henderson, A. F. Izmaylov, G. E. Scuseria and A. Savin, J. Chem. Phys. **127**, 221103 (2007).
- ¹⁶¹ T. M. Henderson, A. F. Izmaylov, G. E. Scuseria and A. Savin, J. Chem. Theory Comput. **4**, 1254 (2008).
- ¹⁶² A. Savin, J. Toulouse and H. J. Flad, Int. J. Quantum Chem. **100**, 1047 (2004).
- ¹⁶³ E. Weintraub, T. M. Henderson and G. E. Scuseria, J. Chem. Theory Comput. **5**, 754 (2009).
- ¹⁶⁴ Jeng-Da Chai and M. Head-Gordon, J. Chem. Phys. **128**, 084106 (2008).
- ¹⁶⁵ A.D. Becke, J. Chem. Phys. **107**, 8544 (1997).
- ¹⁶⁶ T. M. Henderson, B. G. Janesko and G. E. Scuseria, J. Chem. Phys. **128**, 194105 (2008).
- ¹⁶⁷ S. Kozuch and J. M. L. Martin, J. Comp. Chem. **34**, 2327 (2013).
- ¹⁶⁸ S. Grimme, J. Chem. Phys. **124**, 034108 (2006).
- ¹⁶⁹ K. Sharkas, J. Toulouse and A. Savin, J. Chem. Phys. **134**, 064113 (2011).
- ¹⁷⁰ C. Pisani, M. Busso, G. Capecchi, S. Casassa, R. Dovesi, L. Maschio, C. Zicovich-Wilson, and M. Schütz, J. Chem. Phys. **122**, 094133 (2005).
- ¹⁷¹ C. Pisani, L. Maschio, S. Casassa, M. Halo, M. Schütz, and D. Usvyat, J. Comp. Chem. **29**, 2113 (2008).
- ¹⁷² T. Schwabe and S. Grimme, Phys. Chem. Chem. Phys. **8**, 4398 (2006).
- ¹⁷³ A. Karton, A. Tarnopolsky, J. F. Lamere, G. C. Schatz and J. M. L. Martin, J. Phys. Chem. A **112**, 12868 (2007).
- ¹⁷⁴ L. Maschio, J. Chem. Theory Comput. **7**, 2818 (2011).
- ¹⁷⁵ R. Peverati and D. G. Truhlar, Phil. Trans. R. Soc. A, (2014), in press.
- ¹⁷⁶ J. Heyd, J. E. Peralta, G. E. Scuseria and Martin R. L., J. Chem. Phys. **123**, 174101 (2005).
- ¹⁷⁷ P. Hao, Y. Fang, J. Sun, G. I. Csonka, P. H. T. Philipsen, and J. P. Perdew, Phys. Rev. B **85**, 014111 (2012).
- ¹⁷⁸ P. Haas, F. Tran, P. Blaha, and K. Schwarz, Phys. Rev. B **83**, 205117 (2011).
- ¹⁷⁹ O. Madelung, *Semiconductors: Data Handbook* (Springer-Verlag Berlin Heidelberg, New York, 3rd ed., 2004).
- ¹⁸⁰ R. Pässler, Phys. Status Solidi B-Basic Res. **216**, 975 (1999).
- ¹⁸¹ D. M. Roessler and W. C. Walker, Phys. Rev. **166**, 599 (1968).
- ¹⁸² B. Civalleri, D. Presti, R. Dovesi and A. Savin, in *Chemical Modelling: Applications and Theory* (The Royal Society of Chemistry, 2012), vol. 9, p. 168.
- ¹⁸³ M. J. Lucero, T. M. Henderson and G. Scuseria, J. Phys.: Condens. Matter **24**, 145504 (2012).
- ¹⁸⁴ I. J. Bush, S. Tomic, B. G. Searle, G. Mallia, C. L. Bailey, B. Montanari, L. Bernasconi, J. M. Carr, and N. M. Harrison, Proc. R. Soc. A: Math. Phys. Eng. Sci **467**, 2112 (2011).
- ¹⁸⁵ R. Orlando, M. Delle Piane, I. J. Bush, P. Ugliengo, M. Ferrabone, and R. Dovesi, J. Comput. Chem. **33**, 2276 (2012).