PCCP

Dynamic Article Links ▶

www.rsc.org/pccp

PERSPECTIVE

CRYSCOR: a program for the post-Hartree–Fock treatment of periodic systems†

Cesare Pisani,[a] Martin Schütz,*[b] Silvia Casassa,[a] Denis Usvyat,[b] Lorenzo Maschio,[a] Marco Lorenz[b] and Alessandro Erba[a]

CRYSCOR is a periodic *post*-Hartree–Fock program based on local functions in direct space, *i.e.*, Wannier functions and projected atomic orbitals. It uses atom centered Gaussians as basis functions. The Hartree–Fock reference, as well as symmetry information, is provided by the CRYSTAL program. CRYSCOR presently features an efficient and parallel implementation of periodic local second order Møller–Plesset perturbation theory (MP2), which allows us to study 1D-, 2D- and 3D-periodic systems beyond 1000 basis functions per unit cell. Apart from the correlation energy also the MP2 density matrix, and from that the Compton profile, are available. Very recently, a new module for calculating excitonic band gaps at the uncorrelated Configuration-Interaction-Singles (CIS) level has been added. Other advancements include new extrapolation techniques for calculating surface adsorption on semi-infinite solids. In this paper the diverse features and recent advances of the present CRYSCOR version are illustrated by exemplary applications to various systems: the adsorption of an argon monolayer on the MgO (100) surface, the rolling energy of a boron nitride nanoscroll, the relative stability of different aluminosilicates, the inclusion energy of methane in methane–ice-clathrates, and the effect of electron correlation on charge and momentum density of α-quartz. Furthermore, we present some first tentative CIS results for excitonic band gaps of simple 3D-crystals, and their dependence on the diffuseness of the basis set.

[a] *Dipartimento di Chimica IFM, and Centre of Excellence NIS (Nanostructured Interfaces and Surfaces), Università di Torino, via Giuria 5, I-10125 Torino, Italy*
[b] *Institute for Physical and Theoretical Chemistry, Universität Regensburg, Universitätsstrasse 31, D-93040 Regensburg, Germany. E-mail: martin.schuetz@chemie.uni-regensburg.de*
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c2cp23927b

I. Introduction

In the past decade, quantum chemical wavefunction (Wf) based methods started to enter the field of solid state physics and chemistry, a realm dominated almost exclusively by (Kohn–Sham) density functional theory (DFT) with local exchange–correlation functionals. Standard DFT is known

Cesare Pisani

Cesare Pisani became a full professor in Quantum Chemistry at the Torino University in 1981. He is well known for the pioneering works in the field of solid state quantum chemistry. In the seventies he, with his collaborators, developed a first 3D-periodic Hartree–Fock method and implemented it in the CRYSTAL program. Besides this program, which remains one of the most powerful tools for solid state simulations, he developed the EMBED program and the underlying theory to study local defects in crystals. In 2000 he initiated the CRYSCOR project, and since then focused on post-Hartree–Fock methods for periodic systems. Cesare Pisani tragically died in the mountains in July 2011.

Martin Schütz

Martin Schütz is a professor of theoretical chemistry at the University of Regensburg (Germany). He received his PhD in physical chemistry from the University of Berne (Switzerland) in 1993 and spent postdocs 1993–1994, and 1994–1997 at the ETH Zürich (Switzerland), and the University of Lund (Sweden). In 2001 he completed his habilitation in theoretical chemistry at the University of Stuttgart (Germany) and was appointed as a lecturer. In 2004 he accepted a professorship at the University of Regensburg. His research interests are in the area of ab initio electronic structure methods for molecules and solids. He is one of the main authors of the MOLPRO and CRYSCOR programs.

to possess certain deficiencies, like the self-repulsion and self-correlation of electrons (a single electron has a non-vanishing correlation energy), or the lack of long-range van der Waals dispersion (important in any 3D extended system,[1] and totally crucial for molecular crystals). Furthermore, within the DFT method alone there is virtually no systematic way to estimate or improve the quality of the results. The absence of dispersion in DFT can be to some extent circumvented by Grimme's empirical D-corrections.[2,3] However, the empiricism of this approach and the double counting of the correlation energy derogate accuracy and predictive power of the method (see *e.g.* ref. 1, 4 and 5 and also Sections IV B and IV C of this paper). A more rigorous but also computationally much more expensive approach to van der Waals dispersion in the context of DFT is the use of the random phase approximation as the correlation functional.[6]

Wf-based methods do not share the inherent problems of DFT and, moreover, offer a hierarchy of methods to systematically improve the accuracy of a calculation. On the other hand, Wf-based methods are more complicated and thus computationally much more demanding than DFT. In particular, the unfavorable scaling of the computational cost with system size of Wf-based methods is problematic, and, in the context of infinite (2D or 3D periodic) systems, like surface slabs or crystals, even devastating: the simplest post Hartree–Fock method, Møller–Plesset perturbation theory (MP) of second order (MP2), has a scaling of $\mathcal{O}(\mathcal{N}^5)$ with system size $\mathcal{N}$ (*e.g.* the number of electrons), which increases to even $\mathcal{O}(\mathcal{N}^7)$ for coupled cluster singles doubles with perturbative triples correction, CCSD(T), which is the "gold standard" of quantum chemistry. This scaling wall has been overcome in molecular quantum chemistry by exploiting the fact that electron correlation in insulators is a short-range effect. The use of Slater determinants built from *local* and not necessarily mutually orthogonal one-electron functions (orbitals) allows for approximations that exploit the short-range character of electron correlation. Local correlation methods of that sort were already proposed and implemented in the eighties,[7–11] and later it was shown that indeed $\mathcal{O}(\mathcal{N})$ methods can be devised on that basis,[12] even at the level of coupled cluster theory.[13–16] A next step was the

Silvia Casassa

Silvia Casassa is a faculty researcher at the theoretical chemistry group of the Torino University since 2004. She received her PhD from the same university under the supervision of Cesare Pisani. As a postdoc she spent one year at the University of Marseille and six months at the University of Paris VI. She has been developing and applying embedding schemes for localized defects in crystals (she is one of the authors of the EMBED program). Presently, her research activity is focused on the CRYSCOR project.
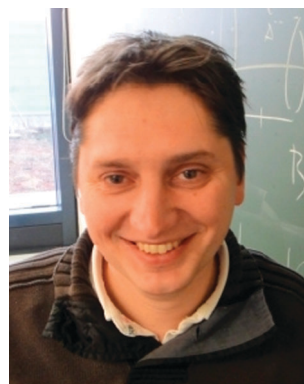
Denis Usvyat

Denis Usvyat is a research fellow at the University of Regensburg. He received his doctoral degree at Saint-Petersburg state University in 2004 (supervisor: Prof. Smirnov). Thereafter, he spent a few months at the University of Stuttgart and since then he has been working in Regensburg (till 2010 as a postdoc of Prof. Schütz). His research is mainly devoted to development of accurate quantum chemical electronic structure techniques for the ground and excited states of solids. Besides that he studies van der Waals interactions in extended systems, and contributes to development of new methods for extended molecules. He is one of the main authors of the CRYSCOR program.

Lorenzo Maschio

Lorenzo Maschio is a faculty researcher in the theoretical chemistry group at the Torino University since 2010. He received his PhD in Materials Science and Technology in the same University in 2006 (supervisor: Cesare Pisani). His research interests are the development and implementation of methods and tools for electronic structure calcula-tions in periodic systems. His additional interests include the application of such methods to the study of properties of crystalline systems. He is one of the main authors of the CRYSCOR program.

Marco Lorenz

Marco Lorenz is a PhD student in the group of Martin Schütz at the University of Regensburg, Germany, where he has also received his diploma in chemistry (2008). The topic of his research is the development of local ab initio methods for calculating optical band gaps in periodic systems. He is one of the authors of the CIS module in the CRYSCOR program.

use of density fitting (DF) in local correlation methods, which lead to very efficient programs for calculating electron correlation effects in extended molecular systems.[17–19]

The first real periodic local correlation method, which was applicable to 3D crystals, was first presented conceptually in 2003[20] and then formally in 2005.[21] This was the start of the CRYSCOR program, driven by the Torino and Regensburg theoretical chemistry groups.[22,23] At that time, the program could calculate the correlation energy of crystals with a small unit cell (like diamond or boron nitride) at the MP2 level.

The reference Hartree–Fock crystalline orbitals, expanded in a basis set of Gaussian-type orbitals (GTO), are provided by a Crystal calculation.[24,25] The occupied and virtual orbital spaces are spanned by localized Wannier functions (WF)[26–28] and projected atomic orbitals (PAO), respectively. The list of WF pairs as well as the pair specific excitation spaces (so-called pair domains) are truncated according to distance criteria in direct space, in analogy to the local correlation methods for molecules. Translational as well as point group symmetry is fully exploited.

It should be mentioned at this point that there have been earlier *cluster based* approaches to tackle 3D-periodic systems in the framework of quantum chemical Wf-based methods, like the incremental scheme proposed by Stoll.[29] The latter has been successfully applied to many different crystals over the years.[30–34] However, such cluster based approaches are not really periodic and thus require a careful and non-trivial embedding to simulate the quasi infinite system.[35] Furthermore, recently, non-local *canonical* MP2 and CCSD methods based on the projector-augmented-wave method and using plane waves have been presented.[36,37] A discussion of these methods in more detail is beyond the scope of the present paper. Here, we focus on the local correlation approach for periodic systems, which features quasi-$\mathcal{O}(\mathcal{N})$ scaling of the computational cost with respect to the unit cell size and therefore allows for a treatment of systems with large unit cells.

The first version of the CRYSCOR program[21] was very inefficient due to the four-index integral transformation, and it turned out that the DF technique constitutes indeed a prerequisite to arrive at an efficient local correlation approach for periodic systems. DF allows for a decomposition of the transformed four-index integrals in terms of three-index quantities. DF for periodic systems turned out to be much more complicated than for the molecular case, and different schemes have been implemented over the years, like the local direct-space fitting, the multipole-corrected reciprocal fitting and the direct-reciprocal-decoupled fitting.[23,38,39] Very recently, also a parallel implementation of the density fitted periodic local MP2 (LMP2) method has been realized.[40] It has been demonstrated that the new program yields reasonable speedups of up to 50 cores, which substantially extends the application range of the method.

In this contribution we highlight the capabilities of the CRYSCOR program by several exemplary applications on physically relevant systems. Section III is devoted to the adsorption problem, for which we have chosen the physisorption of argon on the MgO (100) surface. The scaling of the computational cost with unit cell size within the sequential version of the code is also probed in this section, by considering a MgO 3-layer slab with a progressively expanded unit cell of up to 150 atoms. In Section IV three prototypical examples for LMP2 applications using the parallel CRYSCOR version are presented: here we study the formation energy of a 1D Boron Nitride (BN) nanoscroll, the relative stability of 3D covalent aluminosilicate crystals and the cohesive energy of a molecular clathrate crystal.

Apart from the computation of the second-order correlation energy, CRYSCOR can also evaluate the (orbital unrelaxed) MP2 one-electron *position* density matrix[41] and from that (*via* six-dimensional Fourier transform) the MP2 one-electron *momentum* density matrix, whose diagonal parts are the electron charge density $\rho(r)$ and the electron momentum density $\pi(p)$, respectively. The electron charge density, experimentally accessible through diffraction experiments, faithfully reflects the chemical composition and geometry of the system. Furthermore, Bader's theory of "atoms in molecules"[42] and its application to crystals[43] provide specific information on the chemical features of the system *via* a topological analysis of $\rho(r)$ and its derivatives. Through a 2D integration of $\pi(p)$ over a plane perpendicular to a given crystallographic direction, the directional Compton profiles are obtained, which are directly comparable to the outcome of Compton scattering experiments.[44] With CRYSCOR it is thus possible to calculate momentum space properties at the correlated level.

Note that Kohn–Sham (KS) DFT, footing directly on $\rho(r)$ rather than on the density matrix, has no direct link to momentum space. The KS density matrix is constructed from KS orbitals, which refer to the fictitious system of non-interacting electrons with the only constraint of reproducing the exact $\rho(r)$. The KS density matrix therefore has not much to do with the density matrix of the real system, and KS DFT cannot be expected to provide a reasonable description of the momentum density matrix and its derived properties like the directional Compton profile.[45–49] The topic of calculating the density matrix and related quantities using the CRYSCOR code is covered in Section V.

Another new feature of CRYSCOR is a periodic formulation of the local configuration interaction singles (CIS) method for calculating excitonic band gaps, as a first step towards further and more sophisticated correlated approaches. A first version of the program was limited to 1D periodic systems (polymers),[50] yet

Alessandro Erba is currently a postdoc at the theoretical chemistry group at the Torino University. He received his PhD in Materials Science and Technology, under the supervision of Cesare Pisani, from the same University at the end of 2011. His main research interests focus on ab initio solid state quantum chemistry: symmetry properties of localized crystalline orbitals, one-electron properties from the density matrix of crystals, inclusion of nuclear motion

Alessandro Erba

effect on computed electronic properties, local-correlation methods for periodic systems.

in the mean time it has been generalized to 3D crystals. Some preliminary results for CIS band gaps of simple 3D crystals with focus on basis set effects are presented in Section VI.

In the next section we quickly review the basic aspects of the periodic local post-Hartree–Fock methods. Technical data (geometries, computational parameters, *etc.*) are provided as ESI.†[51]

II. Periodic local post-HF methods

Formalism and implementation of the periodic local MP2 and CIS programs in CRYSCOR have been discussed in detail before[21–23,50,52] and will not be reiterated here. Instead we focus on the present capabilities of the code due to recent advances, which are illustrated by several applications. It is expedient, however, to briefly outline the fundamentals of the method. Both the local MP2 and CIS techniques are essentially based on the local direct space representation of the occupied and virtual spaces. The former is spanned by symmetrized Wannier functions as constructed according to the procedure described in ref. 27 and 28, the latter by projected AOs (PAOs), *i.e.*, AOs, projected onto the virtual space. These are relatively well localized and possess the symmetry of the underlying AOs, but are mutually non-orthogonal and even redundant. Parts of the calculations (*e.g.* the PAO construction[53]) are carried out in reciprocal space when more convenient. Individual objects in direct and reciprocal space representation are connected by Fourier transformation.

The benefits of spanning occupied and virtual spaces by local functions in direct space are two-fold. Firstly, the relevant objects, *i.e.*, integrals, amplitudes (coefficients in the determinantal expansion), *etc.*, are naturally sparse in direct space, and this must be exploited in extended systems to reach efficiency. Secondly, local approximations can be imposed by *a priori* truncating the amplitudes to a much smaller set of local amplitudes considered as relevant on the basis of the short-range nature of dynamic correlation effects. For the LMP2 method this is done in the following way:

(i) The virtual space is truncated according to *WF-pair domains* $\mathscr{D}_{ij}$; all amplitudes related to excitations from the WFs i, j to PAOs *outside* $\mathscr{D}_{ij}$ are zero by construction. $\mathscr{D}_{ij}$ is constructed as the *union* of the two WF domains $\mathscr{D}_i$ and $\mathscr{D}_j$. $\mathscr{D}_i$ comprises the PAOs assigned to centers spatially close to WF i, and can be determined either automatically[54] or manually.

(ii) The WF pair list is truncated according to a distance criterion R_p^{max} applied to the distance R_p between the centers of the WFs in a pair (or between the atoms in the corresponding domains). This *pair approximation* exploits the R_p^{-6} decay behaviour of pair energies in the local representation. Thus, very distant pairs with R_p greater than the cutoff distance R_p^{max} are disregarded when solving the LMP2 equations.[21,22] The pair approximation is vital in the periodic context, since otherwise the number of unknowns in the LMP2 equations becomes infinite. The missing energy contribution of the omitted very distant pairs can, however, be restored *a posteriori* by means of a C_6-extrapolation[22] (*cf.* Section III B).

A further pair approximation concerns the evaluation of the electron repulsion integrals in the WF/PAO basis. These integrals are calculated efficiently *via* periodic density-fitting

techniques.[23,39] However, for pairs with very small overlap of their orbital product distributions it is still much more economical to approximate these integrals by multipole expansion.[22] Therefore, for WF pairs with $R_p > R_{DF}^{max}$ the algorithm switches from density-fitting based integral evaluation to multipole expansion. R_{DF}^{max} must be set to a value beyond which the overlap of the orbital product distributions is negligible.

III. Molecular adsorption on surfaces

One important application area for periodic local correlation methods is molecular adsorption on crystalline surfaces. Representing the periodic surface by means of finite cluster models is difficult, non-unique, and requires sophisticated embedding schemes, especially for adsorption on surfaces of ionic solids.[55,56] Besides, realistic models might require clusters of prohibitively large size. The periodic local MP2 method adequately represents the periodicity of the surface and, at the same time, includes the short- and long-range electron correlation effects on the same footing. The method has been already successfully employed in several applications devoted to adsorption.[57–60]

A. Scaling with unit cell size

The local approximation is of particular importance in the context of adsorption. Often, the adjustment of the adsorbate to the surface geometry implies large supercells, and only the quasi-linear scaling of the computational cost with the number of atoms per cell makes such calculations feasible. Furthermore, since an accurate correlation treatment is much more demanding with respect to (wrt) basis set quality than DFT (especially so for a proper description of van der Waals dispersion), the use of the density fitting technique becomes crucial. It reduces the scaling of the MP2 method wrt basis set size from quartic to cubic, and thus allows for the use of reasonably large basis sets (*e.g.* triple-zeta sets) in conjunction with large unit cells.

In order to demonstrate the scaling behavior of our periodic local MP2 implementation we present results from test calculations performed for the MgO three layer (100) slab with a progressively expanded formal unit cell of up to 150 atoms per cell. A triple-zeta (for Mg) and augmented (d- and f-functions) triple-zeta (for oxygen) basis set, extended WF domains and tightened HF integral screenings parameters[24,61] have been employed in the calculations (see the ESI† for details[51]).

The plots of elapsed times *vs.* unit cell size displayed in Fig. 1 show low (quasi-linear) scaling of the LMP2 and HF calculations. Full linear scaling of the LMP2 part is not expected due to the presence of some higher scaling steps (*e.g.* the PAO construction) with low prefactor. Moreover, for larger supercells some of the buffers can no longer be kept in memory and have to be swapped out on disk, causing I/O overhead and the kinks in the curve. The use of point group symmetry reduces the computational times substantially, particularly so for the HF case.

These supercell calculations also reveal the stability (and size extensivity) of the calculation. Comparing the initial primitive with the largest supercell the HF energy per atom deviates only in the sub-microhartree range. For the MP2 correlation energy the
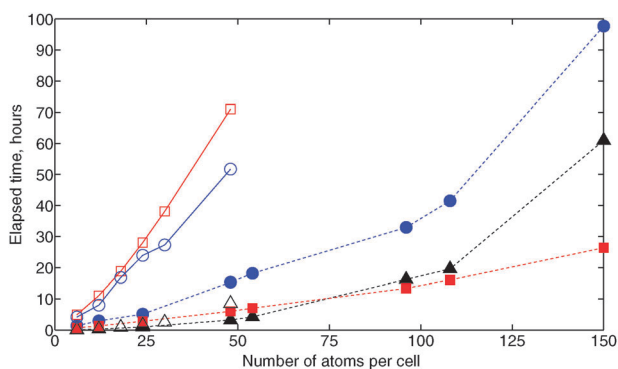
Fig. 1 Elapsed times for HF (red/squares), localization/symmetrization of the WFs (black/triangles) and LMP2 (blue/circles) calculations on a single 8-processor Xeon node. For the calculations the parallel version of Crystal, and sequential version of CRYSCOR (yet with parallel precompiled BLAS libraries) with (filled symbols) and without (open symbols) inclusion of the actual point group symmetry of the slab (16 symmetry operations) has been used.

deviation is still smaller than a few microhartrees per atom, *i.e.*, an order of magnitude less than the density fitting error.[23,39] Probably, these deviations are caused by some small variations in the WFs between the initial primitive, and the supercells (for further discussions see ESI†[51]). In any case, they can safely be disregarded in any application. The memory consumption in the largest supercell calculations is relatively high (up to 20 GByte in the 150 atom case), but is substantially reduced in the parallelized version of the program.[40]

B. Adsorption of argon on the MgO (100) surface

As an example for the adsorption problem we consider the case of the 2×2 square argon monolayer adsorbed on the MgO (100) surface. MgO (100) and the 2×2 argon monolayer are commensurate with respect to each other, but not to the ideal hexagonal argon monolayer, leading to geometrical frustration. The manifold of energetically more favorable and competitive adsorption arrangements of the argon monolayer on MgO (100) (requiring large supercell calculations) was recently studied by some of us; the results of this detailed study will be published elsewhere.[57]

1. Physics of the adsorption process. Fig. 2 shows, as a function of the argon-surface distance, the individual curves of the adsorption energies for argon 2×2 adsorbed on three different sites of MgO (100), *i.e.*, on oxygen, on magnesium, and on the midpoint between two adjacent Mg atoms, respectively (the artificial 2×2 square argon monolayer is constructed such that the individual argon atoms adsorb on equivalent sites of the MgO (100) surface). Due to the sphericity of the argon atoms there is no electrostatic contribution to the adsorption energy. At long range the attraction is governed by van der Waals dispersion. The three potential curves (related to the three adsorption sites) virtually coincide at long range, which exemplifies the isotropy of dispersive interactions. At short range the dominating component is the exponentially growing exchange repulsion. The first-order exchange repulsion is already captured at the HF level. The HF adsorption energy curves, also shown in Fig. 2, do not exhibit any recognizable minima, but at short range their repulsive onsets sensitively
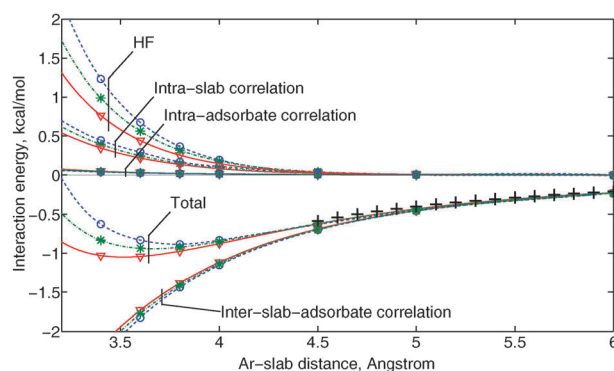

Fig. 2 The interaction energy and its components (HF, intra-slab, inter-slab-adsorbate and intra-adsorbate) between a three-layer MgO (001)-slab and the argon monolayer in the 2×2 supercell arrangement (one argon atom per two Mg atoms on the surface). The on-O (blue/spheres/dashed), on-Mg (red/triangles/solid), and Mg–Mg bridging (green/stars/dash-dotted) adsorption positions are considered. The crosses show the $C_3R^{-3} + C_5R^{-5}$ curve with the C_3 and C_5 parameters obtained in a single-point calculation in the on-oxygen geometry with the slab–argon separation of 3.6 Å.

depend on the individual adsorption site. For the on-O position the repulsive onset starts at larger distance, which is due to the fact that in the MgO slab the electron density is concentrated around the oxygen atoms. Appreciating the isotropy of long range dispersion on the one hand, and the anisotropy of exchange repulsion on the other hand, the most preferable adsorption site is the on-Mg position, as is evident from Fig. 2.

Since local correlation methods employ localized orbitals the correlation part of the adsorption energy can be partitioned into intra-slab, intra-adsorbate, and inter-slab-adsorbate (to be referred to as *inter*-) components (*cf.* Fig. 3).[62,63] Such a decomposition is useful for the analysis of the physics of the adsorption process. Furthermore, the individual components can be scaled to compensate for missing higher-order contributions or basis set deficiencies,[57,64] or, if negligible for the adsorption,
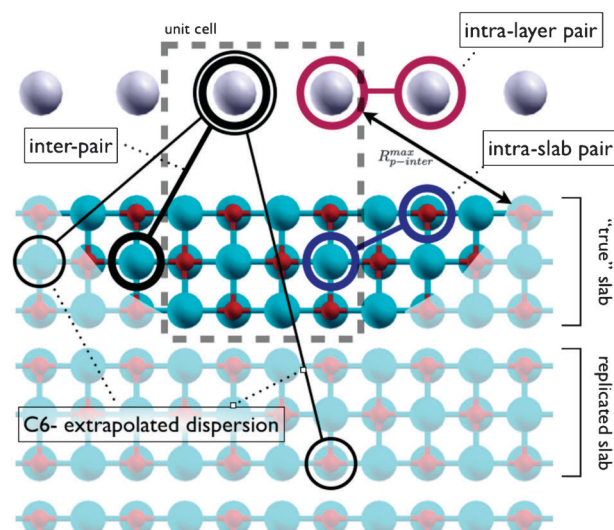

Fig. 3 Schematic representation of the correlation energy partitioning in the local MP2 method and the model of the semi-infinite crystal for the dispersion interaction.

Table 1 The intra-slab correlation energy ($E_{\text{intra-slab}}^{\text{corr}}$) and its contribution ($E_{\text{intra-slab}}^{\text{bind}}$) to binding ($E_{\text{total}}^{\text{bind}}$), depending on the cutoff radius $R_{\text{p-intra-slab}}^{\max}$ for the intra-slab pairs. The amount of the symmetry-irreducible (wrt C_{4v}^1-symmetry group) intra-slab pairs ($N_{\text{intra-slab}}^{\text{pairs}}$) is given, which can be compared to the amount of pairs of other types involved in the calculation, which is 285. The timings include the LMP2 part of the calculations on the full system and the slab alone (with the ghost adsorbate atoms)

$R_{\text{p-intra-slab}}^{\max}/$Å	$N_{\text{intra-slab}}^{\text{pairs}}$	$E_{\text{intra-slab}}^{\text{corr}}/$ hartree	$E_{\text{intra-slab}}^{\text{bind}}/$ kcal mol^{-1}	$E_{\text{total}}^{\text{bind}}/$ kcal mol^{-1}	Elapsed time/hours
0	0	0	0	−1.1030	2.2
2	34	−1.5885	0.2361	−0.8733	4.0
4	110	−1.6876	0.2717	−0.8373	4.8
6	290	−1.6976	0.2779	−0.8316	6.7
8	456	−1.6988	0.2786	−0.8307	7.9
12	904	−1.6992	0.2788	−0.8306	9.0

dropped from the calculation to reduce the computational cost (*vide infra*).

Curves of the intra-slab, intra-adsorbate, and inter-components for argon 2×2 adsorbed on the three different sites of MgO (100) are also displayed in Fig. 2. Evidently, apart from the inter-component comprising long-range dispersion[62,65] also the repulsive intra-slab component is of importance. The latter originates from the reduced ability of slab electrons to mutually correlate due to exchange compression of the electron density caused by the argon layer at short range.

The magnitude of the intra-slab (or intra-adsorbate) components is system-dependent. For some systems it is entirely negligible, which can be exploited to study adsorption on zero or small band gap materials,[58] which by themselves cannot adequately be treated by MP2. But even if the intra-slab component is large (as in the present case) its value quickly converges wrt the pair approximation specified by $R_{\text{p}}^{\max}$. This suggests that a different pair approximation can be used for the inter-pairs $R_{\text{p-inter}}^{\max}$ and intra-pairs $R_{\text{p-intra}}^{\max}$. While the former has to be chosen rather large to capture the relatively slow R^{-6} decay of the dispersion energy with interorbital distance, this is not necessary for the intra-pairs. Table 1 compiles the intra-slab and full correlation interaction energies calculated with different intra-slab pair approximations $R_{\text{p-intra-slab}}^{\max}$.

Evidently, the adsorption energy is already sufficiently accurate for $R_{\text{p-intra-slab}}^{\max} = 4$–6 Å. The computational savings due to such a truncation of the intra-slab pair list are not dramatic for the case of the 3-layer MgO slab with C_{4v}^1-symmetry (a factor of 1.5–2), but for thick low-symmetry slabs the vast majority of the pairs are of intra-slab type, and there such a pair approximation leads to substantial savings, especially concerning memory and disk usage.

2. Extrapolation of the long-range contribution. The intra/inter-pair partitioning can also be used for improving the results of MP2 itself. It is well known that the MP2 method in many cases delivers quite accurate results, but sometimes noticeably overestimates or underestimates the interaction energy. The errors of the method are usually connected to the absence of higher-order dispersion effects.[63,66,67] This deficiency can be corrected by up- or down-scaling the inter-component of the LMP2 interaction energy, with scaling factors obtained from finite-cluster based higher-order correlation calculations with

large basis sets.[57] This allows us to substantially improve the results for cases, where plain periodic local MP2 on the one hand fails to deliver the desired accuracy, while on the other hand a higher-order correlation treatment in the framework of the incremental scheme[29] is either too difficult or computationally too expensive.

The individual LMP2 pair energies beyond a certain WF pair distance R decay as R^{-6}. By analyzing this decay for individual inter-pair energies, the corresponding WF- or atom-pair specific C_6 coefficients can be devised. These can be further utilized to extrapolate the contributions from the WF- or atom-pairs *not explicitly included* in the MP2 calculation. Furthermore, with these coefficients the long-range dispersion between the argon monolayer and the semi-infinite solid (formally constructed by replicating the initial slab at progressively larger inter-slab-adsorbate distances) can approximately be evaluated (*cf.* Fig. 3). The inter-slab-adsorbate (dispersion) energy without and with such an extrapolation technique is displayed in Fig. 4. Without the extrapolation the calculated dispersive interaction is always underestimated, independent of slab thickness. To approach the reference value in this case one needs to increase both the cutoff distance $R_{\text{p-inter}}^{\max}$ and the thickness of the slab. By virtue of the in-slab C_6-extrapolation the converged value can be reached already with moderately large $R_{\text{p-inter}}^{\max}$, yet requiring thick enough slabs and thus a considerable computational effort. But with the slab replication technique the dispersion energy is nearly converged already in a 2-layer slab calculation.

The contributions from the replicated slabs, and the very distant part of the initial slab are in fact quite small. Nevertheless, they are essential for the correct R^{-3} decay behaviour of the van der Waals attraction with slab-adsorbate separation.[68] The atom-pair C_6 (and C_8) coefficients, or the global slab-adsorbate C_3 (and C_5) coefficients, which are computed within one single-point calculation (eventually including a higher-order correction based on finite-cluster calculations, *vide supra*), could also be utilized in the context of less expensive methods like DFT-D or model potentials.[60] Fig. 2 shows that the dispersion energies calculated by means of the C_3 and C_5 coefficients (obtained from an on-O adsorption calculation with a slab–argon distance of 3.6 Å)
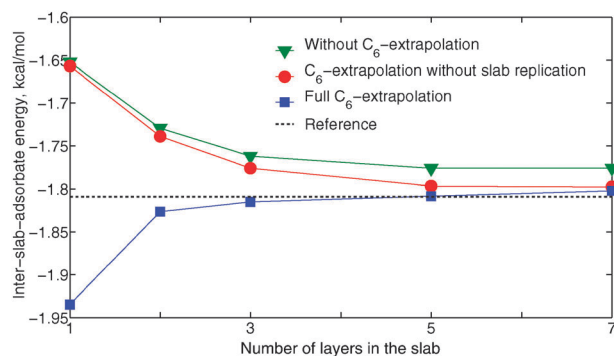

Fig. 4 The inter-component of the interaction energy without (green/triangles) and with the C_6-extrapolated contribution depending on the slab thickness. The latter correspond to the pairs beyond $R_{\text{p-inter}}^{\max}$ but only within the "true" slab (red/circles) or also within the semi-infinite solid (blue/squares), modelled by a replicated slab (*cf.* Fig. 3). The value of 14 Å has been used for $R_{\text{p-inter}}^{\max}$. The reference calculation was done with the 7-layer slab, $R_{\text{p-inter}}^{\max} = 23$ Å and the full C_6-extrapolation. The C_6 coefficients have been automatically fitted to the pair energies in the range from $R_{\text{DF}}^{\max}$ (8 Å) to $R_{\text{p-inter}}^{\max}$.

virtually coincide with long-range MP2 dispersion energies calculated explicitly for various sites and distances.

IV. Parallel periodic local MP2

Recently, a parallel implementation of the periodic local MP2 method was presented,[40] which considerably extends the application range of the method towards larger unit cell sizes and basis sets. This is of particular importance for 3D-crystals, where the dense 3D packing leads to an earlier onset of computational bottlenecks than for 2D-surfaces. Parallelization can indeed be used to eliminate (or at least reduce) computational bottlenecks, but also memory bottlenecks (by recomputing intermediates, which would be computationally too costly in a sequential implementation). The parallel implementation, as discussed in detail in ref. 40, uses exclusively directives of the Message Passing Interface (MPI), while the modifications of the sequential code are kept to a minimum. For the diverse datasets of computational intermediates, which are too large to be kept in memory, efficient disk paging algorithms were devised in order to minimize I/O overhead. On architectures featuring shared filesystems, shared files are used either for all cores involved in the calculation, or for subgroups thereof.

Since the periodic LMP2 code is very heterogeneous with distinct features and peculiarities in different parts, different strategies are adopted for distributing CPU load and memory requirements. The use of distributed memory algebra routines is necessary in those steps, where the size of the matrices involved scales quadratically with the number of atoms per unit cell. In all other parts where the local approximation of the virtual space implies compact local matrices independent of the number of atoms per unit cell, a coarse grain parallelization strategy has mostly been followed, *i.e.*, parallelization according to the index of the atoms in the reference unit cell. Such an approach requires only minor modifications of the sequential code and is conserving its loop structure, but nevertheless is efficient for large unit cell sizes. On the other hand, load imbalance and poor speedups occur for systems with few atoms per unit cell, or with a very different number of basis functions on different atoms.

The parallelization, but also several other technical advancements, substantially improve the efficiency of the periodic local MP2 program, such that presently crystalline systems with a rather large number of basis functions per unit cell can be treated. This opens the door to applications at the cutting-edge of current research. As an illustration of these new capabilities of the code we present here briefly first results from three ongoing studies dedicated to (i) the formation energy of nanoscrolls, (ii) the relative stabilities of different aluminosilicate polymorphs, and (iii) the inclusion energy of methane molecules inside a clathrate cage, which will be reported in detail elsewhere. These three cases are computationally and methodologically quite challenging, for which DFT with standard functionals fails to provide conclusive and authoritative answers.

A. A nanoscroll: from 2D to 1D

As a demonstration for the performance of the new parallel periodic LMP2 implementation we report benchmark calculations
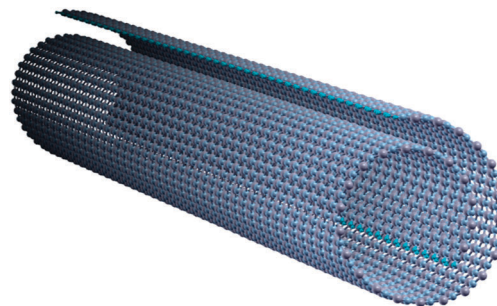

Fig. 5 A portion (45 unit cells are pictured) of the one-dimensionally periodic rolled nano-scroll used in the benchmark calculations of the parallelized CRYSCOR program.

performed for a one-dimensionally periodic boron nitride (BN) nanoscroll. A nanoscroll is a spirally wrapped stripe with a 1D tubular structure resembling that of a multiwalled carbon nanotube (*cf*. Fig. 5), and in fact scrolling is considered as an intermediate stage in the formation of such nanomaterials from colloidal suspensions of layered compounds. Nanoscrolls are expected to inherit some of the excellent properties from both the graphene—or BN—sheet and the related nanotube. On the other hand, such a hybrid structure should also possess some new and unique properties. The interest for such structures directly follows from the huge attention that graphene and derived nanostructures have recently raised in the scientific community.[69]

Here, we compute the "rolling energy" of a BN nanoscroll relative to the related nanostripe (*i.e.* the unrolled system). The atomic positions were optimized at the B3LYP-D2 level. The considered structure has a honeycomb arrangement of B and N atoms, where the two borders of the cut stripe are saturated with hydrogen atoms. The system comprises 122 atoms (60 B, 60 N, and 2 H atoms) per unit cell. For the aug(p,d)-cc-pVDZ basis set[70] used for the MP2 calculation this corresponds to 2274 basis functions per unit cell. Since the diffuse augmented functions cannot be used in the Hartree–Fock procedure due to linear dependencies they are only used at the MP2 stage *via* the dual basis set approach.[53] The number of WF pairs included in the MP2 calculation amounts to more than 70 000. The detailed specifications of the calculation are given in the ESI.†[51]

The rolling energy per BN unit, defined as

$$E^{\text{roll}} = (E^{\text{scroll}} - E^{\text{stripe}})/60 \qquad (1)$$

comprises two main contributions, namely the *repulsive* deformation energy needed to roll up the stripe, and the *attractive* interaction energy of the two adjacent walls now facing each other. The latter originates mainly from van der Waals dispersion (no new chemical bonds are formed), which is not properly described by standard DFT. Calculations performed using a B3LYP functional describe the rolled configuration as not preferable with respect to the unrolled one, with $E^{\text{roll}} = 22.26$ kJ mol^{-1}. The empirical dispersion correction -D2 amounts to -6.21 kJ mol^{-1} only, thus its addition to B3LYP is not sufficient to revert the stability. At the HF level the rolling is not favoured either, with $E^{\text{roll}} = 5.72$ kJ mol^{-1}. However, the MP2 correlation correction is capable of reverting
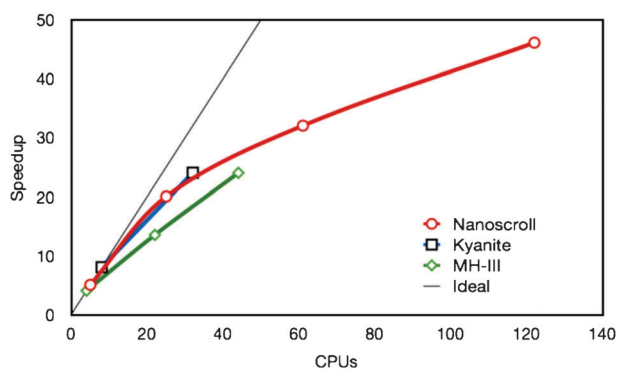
Fig. 6 Parallel performance of some of the calculations presented in Section IV.

Table 2 Relative energy (in kJ mol^{-1} per Al_2O_3 unit) of $\mathscr{A}$ and $\mathscr{S}$ with respect to $\mathscr{K}$. The geometries used in the MP2 calculations were optimized at the PBE0 level

	$\mathscr{A}-\mathscr{K}$	$\mathscr{S}-\mathscr{K}$
HF	−19.17	−22.06
SVWN (LDA)	22.26	36.88
PW91 (GGA)	−5.21	4.88
PBESOL (GGA)	10.32	26.36
B3LYP (hybrid)	−16.50	−13.90
PBE0 (hybrid)	−0.34	10.93
PBE0-D2 (hybrid + disp)	35.76	33.66
HF + MP2	17.66	4.94
Exp.[a]	1.40 (±3.40)	4.22 (±3.43)
Exp.[b]	3.86	7.72

[a] Ref. 72. [b] Ref. 73.

the stability of the two configurations, predicting a rolling energy at the HF + MP2 level of −4.60 kJ mol^{-1}.

Fig. 6 displays the speedups of different parallel LMP2 calculations performed for the BN nanoscroll running on up to 122 cores. The reference calculation was performed on 6 cores. As in previous benchmarks,[40] reasonable speedups are observed for a moderate number of cores (up to 20–30), while the saturation around 60 cores is worse than in other cases. For the first time we have tested here the code on more than 100 cores, which can be considered as the border line between parallel and massively parallel computations. Even though CRYSCOR was not designed for massively parallel computations (*vide supra*), a significant gain is still observed when using up to 122 cores. The overall wall clock time of this calculation is approximately 6 hours, compared to about 10 hours required for a 61 core calculation.

B. Relative energy of 3D covalent crystals: aluminosilicate polymorphs

The relative stability of Al_2SiO_5 ortho-silicate polymorphs (Kyanite, $\mathscr{K}$, Andalusite, $\mathscr{A}$, and Sillimanite, $\mathscr{S}$, whose crystallographic structures are schematically displayed in Fig. 7) is still controversial. On the experimental side, the error bars exceed the small energy differences involved (*cf.* Table 2), while previous quantum chemical calculations based on DFT exhibit an even larger variation of values depending on the choice of the exchange–correlation functional and could not provide a reliable answer to this problem either (*cf.* Table 2
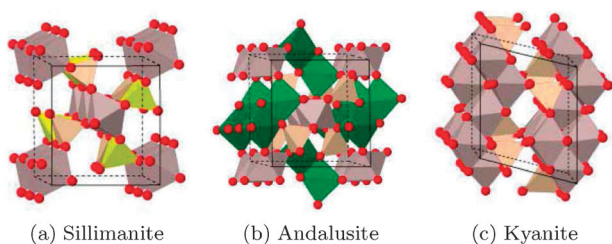

Fig. 7 Crystalline structure of three Al_2SiO_5 polymorphs. Hexa-, penta-, and tetra-coordinated aluminium atoms are represented in gray, green and yellow, respectively (the corresponding coordination polyhedra are reported too); silicon and oxygen tetrahedra are in beige and red, respectively. The picture is published by courtesy of *Int. J. Quantum Chem.*, from ref. 71.

and ref. 71). However, the relative stability of these systems is of great interest in the context of simulations of geological processes. Furthermore, the peculiar structural and chemical features and the size of these systems render the question about their relative stability at the MP2 level an interesting application project and an excellent example for the capabilities of the CRYSCOR program.

For an accurate assessment of the relative stability of these three polymorphs all the contributions to Gibbs's free energy should be considered, that is, the electronic, the pressure–volume, the entropic, and the zero-point terms. Yet it has been demonstrated in ref. 71, that all but the electronic contributions are expected to affect the computed relative stability at most by 2–3 kJ mol^{-1}, an amount that is usually smaller than the experimental uncertainty affecting the measurements.[74] Therefore, solely the electronic contribution is calculated here and used to assess the relative energy of these compounds.

In a first step the geometries of the three polymorphs $\mathscr{K}$, $\mathscr{A}$, and $\mathscr{S}$ were optimized at the DFT level using the PBE0 hybrid functional. This functional yields the best agreement between DFT and experimental structures at 298 K, with the unit cell volume of all the compounds overestimated by only 0.3% with respect to the experimental data.[71] The local MP2 calculations then were performed at these optimized structures. A homogeneous 4-atom *domain* $\mathscr{D}_i$ (the central O plus two Al and one Si atoms) was used for each WF in all polymorphs. For all calculations the 86-311G** basis set on silicon, the 8-411G* on oxygen, and the 86-111G* on aluminium were employed, which comprise 624 contracted Gaussians per cell.

The resulting HF, and HF + MP2 relative energies are compiled in Table 2, along with the corresponding DFT values for different functionals. Experimental values taken from ref. 72 and 73 are also given, for convenience. The local MP2 results establish a clear correlation between the density and the stability of the different phases: the stability order $\mathscr{K} > \mathscr{S} > \mathscr{A}$ coincides with the order of the packing densities. The agreement of the LMP2 relative energies with experiment is quite good, apart from a certain overestimation of the $\mathscr{A}-\mathscr{K}$ stability. The DFT results, on the other hand, wildly scatter for the different functionals (in line with previous results),[71] and are nowhere near neither the LMP2, nor the experimental value. Of those tried, the PBESOL functional provides the best estimate. Addition of Grimme's dispersion correction -D2[2] to the PBE0 functional, followed by geometry optimization of all the
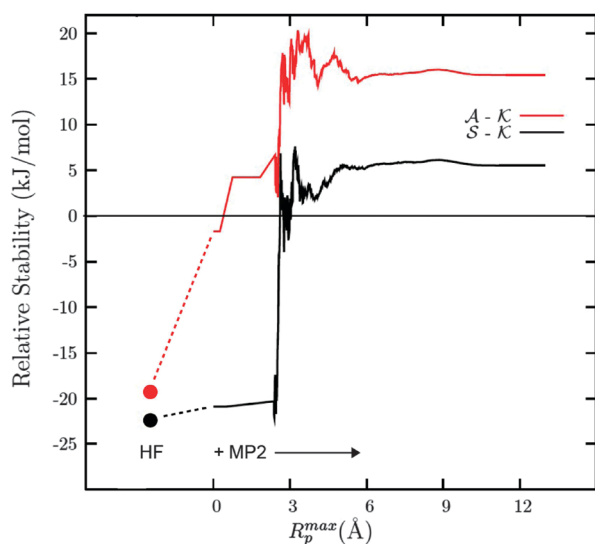
7622 | *Phys. Chem. Chem. Phys.*, 2012, **14**, 7615–7628

This journal is © the Owner Societies 2012

Fig. 8 HF + MP2 relative stability (in kJ mol^{-1} per Al$_2$O$_3$ unit) of $\mathscr{S}$ (red) and $\mathscr{S}$ (black) with respect to $\mathscr{K}$ as a function of the correlation distance between WFs. The HF relative stabilities are reported as well, as full circles.

structures, changes the order predicted by this Hamiltonian: $\mathscr{K}$ correctly becomes the most stable phase with $\mathscr{A}$ and $\mathscr{S}$ being almost isoenergetic. The PBE0-D2 values however are far off the experimental ones, indicating an overestimation of the dispersive contribution in the denser phase, $\mathscr{K}$, and hence a need for a different calibration of the empirical parameters of the -D2 correction for dense crystalline materials. Furthermore, all the PBE0-D2 structures feature a 1–2% smaller volume relative to the experimentally observed ones, which is another indication of the overestimation of the dispersive contribution by PBE0-D2.

Fig. 8 shows the relative stability of $\mathscr{A}$, and $\mathscr{S}$, vs. $\mathscr{K}$, calculated with different pair approximation $R_p^{\max}$. This dependence indicates whether the short-range or long range (dispersion) or both are responsible for the relative stability between the compounds. Apparently, at the uncorrelated HF level ($R_p^{\max} < 0$) both $\mathscr{A}$, and $\mathscr{S}$ are more stable than $\mathscr{K}$, while the $\mathscr{K}$ phase is becoming most stable when correlation is included. Interestingly, very short-range correlation involving WF pairs having at least one center in common, i.e., $R_p^{\max} = 0$, is almost sufficient to change the relative energy between $\mathscr{A}$ and $\mathscr{K}$, which is not the case for $\mathscr{S}$ vs. $\mathscr{K}$. Increasing $R_p^{\max}$ beyond 3–4 Å, i.e. adding a substantial portion of dispersion, yet is sufficient to make $\mathscr{K}$ the overall most stable phase.

In order to assess the dependence of the results on the crystal geometry, additional MP2 calculations have been performed at the B3LYP optimized geometries (B3LYP overestimates the volume by about 2.5%). At this alternative geometry the MP2 relative energy of $\mathscr{A}$ and $\mathscr{S}$ remains virtually the same as that at the PBE0 geometry. To summarize, $\mathscr{K}$ is found to be the most stable structure, in agreement with experiment. Yet, in contrast to the experiment, $\mathscr{S}$ is, at the MP2 level, significantly more stable than $\mathscr{A}$. This remaining discrepancy might be due to the neglect of other but the electronic terms in the expression of Gibbs's free energy, or due to deficiencies of the basis set. An intrinsic problem of atom centered basis sets,

affecting both DFT and MP2, is the geometry dependence of the basis, as manifested in the basis set superposition error (BSSE). Even though the BSSE of local correlation methods is considerably smaller than that of canonical correlation methods, it is still of the order of the DFT BSSE and thus can lead to an artificial favouring for the more densely packed structures. BSSE free relative energies are obtained by comparing instead of the total energies the counterpoise corrected cohesive energies of the competing polymorphs, as done e.g. in ref. 1. All these effects need to be further explored.

C. van der Waals binding in solids: inclusion energy of CH$_4$ in methane–ice-clathrates

Molecular crystals constitute a very interesting class of materials under many scientific and technological aspects. Over the years a number of correlated finite cluster approaches (sometimes combined with partial periodic treatment, e.g. periodic HF) for such systems have been formulated and applied.[34,75–79] Studies on molecular crystals have already been carried out also in a purely periodic format using earlier versions of the CRYSCOR parallel code, including pressure-induced phase transitions[80,81] and cohesive energies of molecular crystals.[82–84]

A clathrate or cage compound is a molecular crystal representing a lattice of molecules of a first type, which traps molecules of a second type. A clathrate hydrate, in particular, is a special type of gas hydrate, in which a lattice of water molecules encloses molecules of a trapped gas. Clathrate hydrates containing methane are of particular interest. Large amounts of such methane–ice, which is considered as an important potential energy resource of the future, have been discovered under the ocean sea-bed. Deposits seem to be abundant in the continental shelf, contributing also significantly to the stability of the slopes of the shelf. For example, the Storegga slides, which are amongst the largest known landslides causing huge tsunamis, are assumed to be triggered by a decomposition of methane–ice. Furthermore, in astrochemistry, methane hydrates are conjectured to the dominant methane-containing phase in the nebulae from which Saturn, Uranus, Neptune, and their major moons were formed.[85] So, methane hydrates, and in particular their stability, certainly are of great interest, presently. The particular system studied here is classified under the name of MH-III (cf. Fig. 9), and its existence has been discovered only recently.[86] It appears to be stable in a broader range of pressure–temperature values, and to be richer in methane, than other methane hydrate phases.

We have calculated the average *inclusion energy* per CH$_4$ molecule, E_{incl}, of the methane gas in the ice matrix, as

$$E_{\text{incl}} = [E_{\text{MH-III}} - 4E_{\text{M}} - E_{\text{H-III}}]/4 \qquad (2)$$

where E_{M} is the energy of an isolated methane molecule and $E_{\text{H-III}}$ is the energy of the empty hydrate structure. The experimental geometry of the MH-III phase has been used in all cases. E_{incl} also does not contain any geometrical relaxation energies of the individual fragments. To take into account BSSE effects, ghost functions from CH$_4$ molecules have been retained in the empty clathrate cage, and in a sphere of 4 Å around the isolated methane molecule. Local excitation
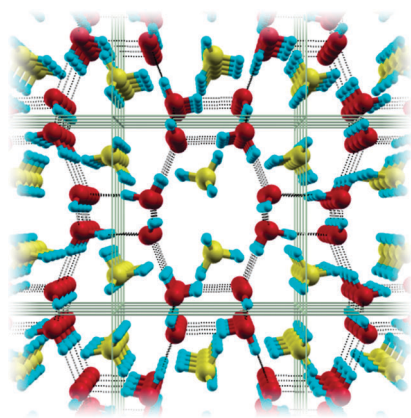
Fig. 9 Crystalline structure of the methane hydrate-III phase.

Table 3 Energies (in kJ mol^{-1}) for the inclusion of methane molecules in the MH-III structure, as defined in eqn (2). The values from HF, MP2 and several DFT approaches are reported

	HF	MP2	B3LYP	B3LYP-D2	PWGGA
E^{incl}	46.48	−7.56	39.60	−22.02	26.67

domains have been restricted, for each WF, to the atoms forming the molecule they belong to.

An unmodified standard molecular cc-pVTZ basis set was employed, with the addition of augmented d- and f-orbitals[87] in the MP2 calculation (the latter by means of the dual basis set technique),[53] leading to a respectable size of 1112 contracted basis functions per unit cell. Table 3 compiles the counterpoise corrected[88] *inclusion energies* obtained with HF, HF + MP2, and DFT.

Of the DFT functionals, evidently, only the dispersion corrected B3LYP-D2 leads to a physically reasonable, attractive inclusion energy, although significantly overestimated in comparison to MP2. Again we see the need for a reparametrization of the empirical parameters of the -D2 correction for crystalline materials.

The parallel speedup for the LMP2 calculation on the MH-III structure is reported in Fig. 6 with respect to a reference calculation on 4 cores. The performance is in this case not excellent, due to load imbalancing issues originating from the very different number of basis functions on C and O atoms, relative to H. On 44 cores, the total wall clock time amounts to less than 3 hours.

V. MP2 density matrix and related properties

All the relevant one-electron properties related to spin-free operators can be calculated from the spin-free *position* one-electron density matrix (DM) $R(r;r')$.[89,90] The electron charge density (ECD) of the system is simply the diagonal part of the DM, $\rho(r) = R(r;r)$. The three-dimensional Fourier Transform (FT) of $\rho(r)$ yields the *form factor* $F(p)$, which in turn can be represented by the discrete (but infinite) set of *structure factors* (SF) as $F(p) \equiv \{F_{hkl}\}$, where h, k, l are Miller's indices labeling crystallographic directions. These structure factors are the primary observables in direct space, which can be measured with X-rays or synchrotron radiation diffraction

experiments. An "experimental" ECD can then be reconstructed *via* the so-called *multipolar models*.[91]

The six-dimensional FT of $R(r;r')$ yields the *momentum DM* $P(p;p')$; its diagonal part $\pi(p) = P(p;p)$ is the electron momentum density (EMD) of the system. The three-dimensional FT of the latter is the *reciprocal form factor* $B(r)$ (also known as *auto-correlation function* or *internally folded density*).[92,93] For each [hkl] crystallographic direction, identified by the unit vector e_{hkl}, two interrelated functions of a single variable can be defined, namely the directional auto-correlation function, $B_{hkl}(r) = B(re_{hkl})$, and the directional Compton profile (CP), $J_{hkl}(p)$ (the latter is simply the one-dimensional FT of $B_{hkl}(r)$). Within the sudden-impulse approximation, $J_{hkl}(p)$ is directly comparable to the outcome of Compton scattering experiments,[44] after correcting these for limited resolution and multiple scattering effects. The "experimental" EMD of the system, even though not uniquely,[94] can then be reconstructed from a sufficiently large set of such directional CPs.[95]

The CRYSCOR program can presently calculate the orbital-unrelaxed MP2 density matrix related to energy derivatives of the MP2 energy with respect to external perturbations.[41,96] This allows for a post-HF treatment for solid state properties such as the ECD (along with a topological analysis according to Bader[42]), SFs, EMDs, CPs, *etc*. In particular, the availability of correlated momentum-space properties is of relevance, since KS-DFT cannot provide a reliable description of such quantities.[45–48,97]

As an example of this feature of CRYSCOR, we discuss here two complementary aspects of the DM of α-quartz as reflected in its EMD and its ECD, and the effect of electron correlation on these. All the computational parameters used in these calculations are identical to those of ref. 49.

Let us first consider its EMD. A simple scheme for analyzing the total EMD of a system is that of considering its anisotropy $\Delta\pi(p)$ with respect to the *spherical average* (SA) function $\bar\pi_{SA}(p)$, $\Delta\pi(p) = \pi(p) - \bar\pi_{SA}(p)$, with $p = |p|$. The left panel of Fig. 10 displays the $\Delta\pi(p)$ quantity calculated at the Hartree–Fock level for momentum vectors p confined to a plane containing the [100] and [001] directions. Evidently, the EMD of α-quartz exhibits a certain anisotropy; the region of maximum anisotropy lies at $|p|$ values between 1.0 and 1.3 a.u. near the [101] direction.

The right panel of Fig. 10 shows the MP2 contribution $\pi^{MP2}(p)$ to the total HF + MP2 EMD. $\pi^{MP2}(p)$ is negative at low momenta ($p < 1.8$) and positive at higher momenta. The effect of electron correlation is that of increasing the kinetic energy of the system by "replacing" electrons with lower momenta by electrons with higher momenta. This can be understood on the basis of the virial theorem. Since electron correlation always leads to a lower (more negative) total energy compared to the mean field theory, the kinetic energy of the correlated system must also be higher. Furthermore, it can be seen that $\pi^{MP2}(p)$ has a minimum around the [101] direction where $\Delta\pi^{HF}(p)$ exhibits a maximum, such that electron correlation also *decreases* the anisotropy of the EMD with respect to the HF reference.

One of the intriguing aspects of α-quartz is its degree of ionicity. Complete ionicity corresponds to net charges of $+4$ and -2 on the silicon and oxygen atoms, respectively. The simple Mulliken population analysis[98] employing DFT/PBE

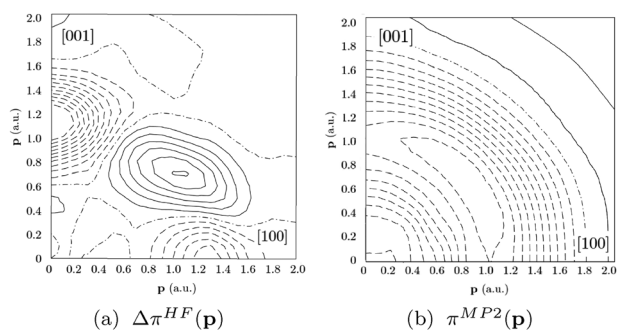(a) $\Delta\pi^{HF}(\mathbf{p})$ (b) $\pi^{MP2}(\mathbf{p})$

Fig. 10 Left panel (a): anisotropy map of the EMD of α-quartz as computed at the HF level in a vertical plane; the step between isolines is 0.01 a.u. and the maximum and minimum values reported are 0.1 and −0.1 a.u.; right panel (b): the MP2 contribution to the total HF + MP2 EMD in the same plane; the step between isolines is 0.002 a.u.; positive and negative regions are represented by continuous and dashed lines, respectively.

Kohn–Sham orbitals yields charges of $+1.66$ and $−0.83$ (other functionals yield similar values),[99] indicating that the ionicity is not as large. Hartree–Fock, on the other hand, yields considerably larger values of $+2.26$ and $−1.13$. Judging from the comparison of DFT and Hartree–Fock, electron correlation appears to decrease the degree of ionicity in α-quartz. In order to investigate this conjecture we consider the MP2 contribution $\rho^{MP2}(\mathbf{r})$ to the ECD of the system.

Fig. 11(a) displays the density difference $\Delta\rho(\mathbf{r})$ between the Hartree–Fock ECD of the real system and the corresponding superposition of non-interacting atomic densities, for $\mathbf{r}$ restricted to a plane containing the two Si–O bonds of one of the three equivalent oxygens. Evidently, the effect of the crystal packing (relative to the isolated and neutral atoms) is that of displacing electron density from silicon to oxygen atoms. Fig. 11(b) shows the related MP2 contribution $\rho^{MP2}(\mathbf{r})$ to the total HF + MP2 ECD. Clearly, electron correlation shifts electron density from oxygen to silicon, hence it *decreases* the degree of ionicity relative to Hartree–Fock, as already anticipated. However, the effect is rather small, the related Mulliken charges are only $−0.02$ and $+0.01$ for silicon and oxygen, respectively.

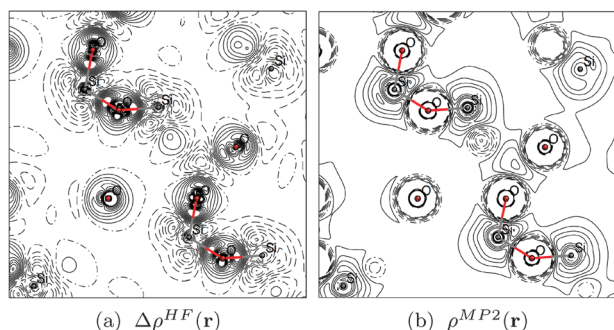(a) $\Delta\rho^{HF}(\mathbf{r})$ (b) $\rho^{MP2}(\mathbf{r})$

Fig. 11 Left panel (a): difference $\Delta\rho(\mathbf{r})$ between the total ECD $\rho(\mathbf{r})$ of α-quartz and the superpositions of non-interacting atomic densities $\rho^{at}(\mathbf{r})$, as computed at the HF level. The separation between isolines is 0.005 |e| bohr^{-3}; right panel (b): the MP2 contribution to the total HF + MP2 ECD of α-quartz. The separation between isolines is 0.0002 |e| bohr^{-3}. Positive and negative regions are represented by continuous and dashed lines, respectively.

VI. Excitonic band gaps of simple 3D crystals

A very recent extension of CRYSCOR is an implementation of the periodic local configuration interaction singles (CIS) method for calculating excitonic band gaps.[50] The CIS exciton wavefunction with the symmetry $\mathbf{k}_{exc}$ is defined as

$$|\Psi_{exc}(\mathbf{k}_{exc})\rangle = \hat{C}_1(\mathbf{k}_{exc})|0\rangle, \text{ with}$$

$$\hat{C}_1(\mathbf{k}_{exc}) = \sum_{ilaA} c_{aA}^{il}(\mathbf{k}_{exc})\hat{E}_{il}^{aA}$$

$$= \sum_{\bar{i}\mathbf{k}_i \bar{a}\mathbf{k}_a} c_{\bar{a}\mathbf{k}_a}^{\bar{i}\mathbf{k}_i}(\mathbf{k}_{exc})\hat{E}_{\bar{i}\mathbf{k}_i}^{\bar{a}\mathbf{k}_a}, \quad (3)$$

in the direct and reciprocal (canonical) space, respectively. Here $c_{aA}^{iI}(\mathbf{k}_{exc})$ denotes the CIS coefficients in the direct space (the orbital and cell indices for WFs and PAOs are i, I and a, A, respectively), while $c_{\bar{a}\mathbf{k}_a}^{\bar{i}\mathbf{k}_i}(\mathbf{k}_{exc})$ are their canonical counterparts ($\bar{i}$, $\bar{a}$ denote the occupied and virtual Bloch orbitals with translational symmetries $\mathbf{k}_i$ and $\mathbf{k}_a$). The spin-conserving one particle excitation operators $\hat{E}_{iI}^{aA}$ (direct space) or $\hat{E}_{\bar{i}\mathbf{k}_i}^{\bar{a}\mathbf{k}_a}$ (reciprocal space) act on the Hartree–Fock reference $|0\rangle$. Presently, only Γ-point ($\mathbf{k}_{exc} = 0$) excitons can be calculated, which correspond to *vertical* excitations ($\mathbf{k}_a = \mathbf{k}_i$) in the canonical basis or periodic excitations $c_{aA}^{iI} = c_{aA\ominus I}^{i0}$ in the direct space. Non-vertical excitations ($\mathbf{k}_{exc} \neq 0$) can presently be calculated by the supercell method.

The CIS eigenvalue problem $\hat{H}_N\hat{C}_1|0\rangle = \omega_{CIS}\hat{C}_1|0\rangle$ is solved in the canonical basis by employing the Davidson diagonalization method.[100] The electron repulsion integrals in $\hat{H}_N\hat{C}_1|0\rangle$ are decomposed in the direct space employing density fitting (DF).[50] The DF integrals and coefficients are evaluated in the local basis of WFs and PAOs and at each iteration contracted with the direct-space images of the trial CIS coefficients. The actual 4-index electron repulsion integrals are thus never assembled, which reduces the nominal scaling of the method with system size to $\mathcal{O}(\mathcal{N}^4)$. The resulting two-electron part of $\hat{H}_N\hat{C}_1|0\rangle$ is then Fourier-transformed to the reciprocal space in each iteration of the Davidson diagonalization procedure.

Our first implementation, reported in ref. 50, was limited to polymers. The reason was that the lattice sums required for the construction of the Coulomb two-electron part of $\hat{H}_N\hat{C}_1|0\rangle$ converge only in the 1D case. Yet in the present version of the program this restriction has been overcome by implementing a multipole expansion for non-overlapping DF integrals in combination with an Ewald-like procedure to carry out the infinite lattice sums in the DF procedure. This new approach will be discussed in detail in a forthcoming article.[52]

In the following we will present some first test calculations of CIS excitonic band gaps of the MgO and LiF crystals and explore how the band gaps depend on the diffuseness of the basis set. The use of atom-centered basis sets in periodic systems is generally a long-standing issue,[61] which does not yet have an ultimate solution. Standard Gaussian basis sets (optimized for molecular calculations) are hardly applicable even at the HF level because of quasi-linear dependencies, leading to numerical instability. It is usually argued, however, that in well-packed periodic systems devoid of voids, diffuse atomic orbitals are needless and even counterproductive.

An accurate description of electron correlation in the ground state of periodic systems requires somewhat more diffuse basis functions compared to HF, but still substantially less diffuse than those of molecular basis sets.[101] The situation, though, is different for excitation energies, and the basis set problem is much more delicate even for molecules. Since standard Gaussian basis sets are usually optimized for the description of electronic ground states they naturally are less optimal for the excited states. Due to this lack of balance, the excitation energies as the difference of ground and excited state energies are generally overestimated. In molecular calculations, augmenting the basis set by additional diffuse functions usually leads to an appreciable lowering of the excitation energy (of the same valence state). In the following we investigate the effect of increasing diffuseness of the basis set on the fundamental and excitonic band gaps of dense 3D crystals.

First, we consider the excitonic (optical) Γ-point band gap of the MgO crystal. The basis sets we used for the calculations are 8-511G* for Mg and 8-51G* for oxygen, which have been employed in a number of periodic HF or DFT calculations.[102,103] It is noteworthy that the most diffuse Mg sp-primitive Gaussians of that basis are, with an exponent of 0.28 a.u., far less diffuse than those of any molecular Mg basis set. For instance, in Pople's 3-21G basis set the most diffuse primitive has an exponent of 0.0464011 a.u. Such a diffuse basis set is definitely not applicable in periodic calculations, but even much less diffuse ones, which can be processed in the periodic SCF procedure apparently do not improve the ground state energy. Yet this is no longer fulfilled for the excited states. Fig. 12(a) shows the dependence of the CIS band gap, the HF total energy, and the DFT (B3LYP and LDA) Γ-point difference between the highest occupied and lowest unoccupied crystalline orbitals (HOCO–LUCO) on the value of the exponent of the most diffuse orbital. For the largest value of the exponent (0.28 a.u.) the CIS band gap is much closer to the experimental value than the HF HOCO–LUCO difference, but still grossly overestimates it (by about 5 eV). However, the diffuseness of the sp-primitives has a strong effect on both the CIS energy and the HF HOCO–LUCO difference: decreasing the exponent of the sp-primitives from 0.28 a.u. to 0.14 a.u. decreases both quantities by nearly 2 eV. The curve of the CIS energy, for

which the HF HOCO–LUCO difference is its zeroth-order contribution, is nearly parallel to the curve of the latter. Evidently, the lowering of the band gap indeed is due to an improved description of the excited state, rather than a degradation in the ground state energy: the latter remains virtually unaffected by the variation of the basis set. Still, very diffuse functions (like in molecular basis sets) are still not of much use in crystals even for excited states: CIS energy and HF HOCO–LUCO difference both saturate for exponents smaller than 0.15 a.u.

Interestingly, the DFT HOCO–LUCO difference, which is often interpreted and used as the optical band gap (for non-hybrid functionals the time-dependent DFT band gap in solids indeed coincides with the HOCO–LUCO difference[104,105]), is also severely affected by the diffuseness of the basis set. For instance, the B3LYP HOCO–LUCO difference obtained for the largest exponent is indeed quite close to the experimental value. However, with improvement of the basis set it decreases parallel to HF, and for exponents smaller than 0.15 a.u. it finally deviates noticeably from the experiment.

In the case of the LiF crystal (shown in Fig. 12(b)) the decrease of the exponent of the most diffuse lithium sp-shell (from 0.22 to 0.14 a.u.) does not alter the value of the band gap much. At the same time, the initial CIS error in this case is much smaller than in the MgO case (around 2.5 eV).

To summarize, these preliminary calculations indicate that basis set effects can play an important role in excited states in 3D solids even at the uncorrelated CIS level. With well balanced basis sets the CIS error can be reduced to 2–3 eV, or even less, which is about twice as large as the usual CIS error in molecular calculations (1–2 eV). The accuracy can be further improved by including dynamic correlation (or screening) effects, which are larger in 3D solids than in 2D slabs or molecules.[106]

VII. Summary and conclusions

In this paper we report on recent advances in the CRYSCOR program. The present version features an efficient and parallel implementation of periodic local MP2, which allows calculations of correlated cohesive energies for 3D-periodic systems with large unit cells in decent basis sets. As exemplary applications the "rolling energy" of a boron nitride nanoscroll, the relative stability of aluminosilicate polymorphs, and the inclusion energy of a methane–ice-clathrate are discussed. It turns out that in many cases density functional theory with standard functionals provides unsatisfactory results, even if the lack of van der Waals dispersion is corrected for e.g. by the popular empirical -D correction proposed by Grimme. A description of periodic systems at the MP2 level (and beyond) is thus highly desirable.

A particularly interesting application area for the periodic local MP2 method is molecular adsorption on surfaces. Here, a reasonable description of van der Waals dispersion, often constituting the major fraction of the adsorption energy, is particularly important. Moreover, due to geometrical frustration of the adsorbate layers non-commensurate to the surface, large supercells need to be used, and therefore a favorable scaling of the computational cost with unit cell size is mandatory. In the present work we consider the 2 × 2 argon monolayer adsorbed
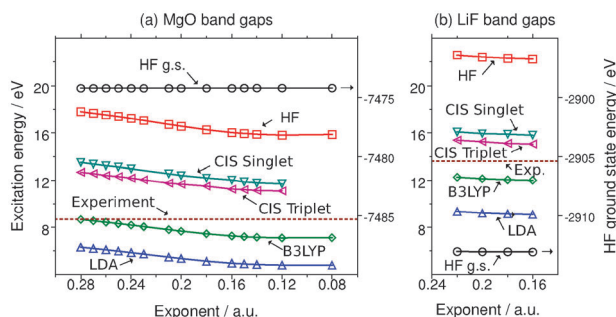

Fig. 12 The HF ground state energies (in the right y-axis) and Γ-point HOCO–LUCO (HF, DFT) and CIS band gap energies (in the left y-axis) for the MgO (a) and LiF (b) crystals as functions of the exponent of the most diffuse sp-shell in the Mg or Li basis set, respectively. The widths of the plots have been chosen to maintain the same scale of the x-axis.

on the MgO (100) surface. Quasi-linear scaling of the computational cost with unit cell size is demonstrated for a set of calculations on progressively larger MgO (100) supercells.

A scheme for adjusting the inter-slab-adsorbate component of the LMP2 correlation energy to higher-order methods (on the basis of finite cluster calculations) is possible, which compensates to large extent the inherent under- or over-estimation of MP2 adsorption energies. Finally, a technique is proposed to directly calculate Wannier function- and atom-specific van der Waals coefficients, and to evaluate with these the van der Waals attraction between the mono-layer and the *semi-infinite solid*. Utilizing this extrapolation technique very quick convergence of the adsorption energy with slab thickness is observed; already a two-layer slab provides sufficient accuracy, and already with a three-layer slab convergence to the semi-infinite solid is reached.

A further extension of the CRYSCOR program is a module for calculating the MP2 density matrix. From the latter the correlated electron momentum density matrix, and from this the directional Compton profiles can be obtained. As an example, the effect of electron correlation on the anisotropy of the EMD of α-quartz is investigated. As anticipated on the basis of the virial theorem, electron correlation increases the kinetic energy of the system. Electron density is shifted from lower to higher momenta. Furthermore, it is observed that a local minimum of the incremental MP2 momentum density meets a maximum of the Hartree–Fock density. Thus, a further effect of electron correlation in α-quartz is to reduce the anisotropy of the uncorrelated electron momentum density.

Very recently, a periodic local configuration interaction singles (CIS) module for calculating excitonic band gaps has been added to CRYSCOR. As for the MP2 occupied and virtual orbital spaces are spanned by Wannier functions, and PAOs, respectively. The integrals, and their contractions, are evaluated in direct space, while the Davidson diagonalization is carried out in reciprocal space. Here, we present first tentative calculations for some simple 3D-crystals. To our knowledge, these are the first pure CIS calculations carried out for 3D-periodic systems. Later work will be directed towards (i) a hybrid approach combining long-range CIS with short-range TD-DFT by splitting of the Coulomb operator, and (ii) including electron correlation effects at the level of second-order perturbation theory. On the other hand, the problem of basis set imbalance in the description of electronic ground and excited states asks for a rigid solution, particularly in the context of periodic systems. Since diffuse functions are of prime importance here, augmenting the standard Gaussian basis set by low energy plane waves might be a promising way forward.

Acknowledgements

References

1 M. Halo, C. Pisani, L. Maschio, S. Casassa, M. Schütz and D. Usvyat, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**, 035117.
2 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787.
3 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
4 B. Civalleri, C. M. Zicovich-Wilson, L. Valenzano and P. Ugliengo, *CrystEngComm*, 2008, **10**, 405.
5 S. Tosoni and J. Sauer, *Phys. Chem. Chem. Phys.*, 2010, **12**, 14330.
6 J. Harl and G. Kresse, *Phys. Rev. Lett.*, 2009, **103**, 056401.
7 P. Pulay, *Chem. Phys. Lett.*, 1983, **100**, 151.
8 S. Saebø and P. Pulay, *Chem. Phys. Lett.*, 1985, **113**, 13.
9 P. Pulay and S. Saebø, *Theor. Chim. Acta*, 1986, **69**, 357.
10 S. Saebø and P. Pulay, *J. Chem. Phys.*, 1987, **86**, 914.
11 S. Saebø and P. Pulay, *J. Chem. Phys.*, 1988, **88**, 1884.
12 M. Schütz, G. Hetzer and H.-J. Werner, *J. Chem. Phys.*, 1999, **111**, 5691.
13 M. Schütz, *J. Chem. Phys.*, 2000, **113**, 9986.
14 M. Schütz and H.-J. Werner, *J. Chem. Phys.*, 2001, **114**, 661.
15 M. Schütz, *J. Chem. Phys.*, 2002, **116**, 8772.
16 M. Schütz, *Phys. Chem. Chem. Phys.*, 2002, **4**, 3941.
17 H.-J. Werner, F. R. Manby and P. Knowles, *J. Chem. Phys.*, 2003, **118**, 8149.
18 M. Schütz, H.-J. Werner, R. Lindh and F. R. Manby, *J. Chem. Phys.*, 2004, **121**, 737.
19 M. Schütz and F. R. Manby, *Phys. Chem. Chem. Phys.*, 2003, **5**, 3349.
20 C. Pisani, *THEOCHEM*, 2003, **621**, 141.
21 C. Pisani, M. Busso, G. Capecchi, S. Casassa, R. Dovesi, L. Maschio, C. Zicovich-Wilson and M. Schütz, *J. Chem. Phys.*, 2005, **122**, 094133.
22 C. Pisani, L. Maschio, S. Casassa, M. Halo, M. Schütz and D. Usvyat, *J. Comput. Chem.*, 2008, **29**, 2113.
23 M. Schütz, D. Usvyat, M. Lorenz, C. Pisani, L. Maschio, S. Casassa and M. Halo, in *Accurate Condensed Phase Quantum Chemistry*, ed. F. R. Manby, CRC Press, Boca Raton, FL, 2010, p. 29.
24 R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, K. Doll, N. M. Harrison, B. Civalleri and I. J. Bush, *et al.*, *CRYSTAL09 User's Manual*, Università di Torino, Torino, 2010, http://www.crystal.unito.it.
25 R. Dovesi, R. Orlando, B. Civalleri, C. Roetti, V. R. Saunders and C. M. Zicovich-Wilson, *Z. Kristallogr.*, 2005, **220**, 571.
26 G. H. Wannier, *Phys. Rev.*, 1937, **52**, 191.
27 C. M. Zicovich-Wilson, R. Dovesi and V. R. Saunders, *J. Chem. Phys.*, 2001, **115**, 9708.
28 S. Casassa, C. M. Zicovich-Wilson and C. Pisani, *Theor. Chem. Acc.*, 2006, **116**, 726.
29 H. Stoll, *Phys. Rev. B: Condens. Matter*, 1992, **46**, 6700.
30 K. Rościszewski, B. Paulus, P. Fulde and H. Stoll, *Phys. Rev. B: Condens. Matter*, 1999, **60**, 7905.
31 K. Rosciszewski, K. Doll, B. Paulus, P. Fulde and H. Stoll, *Phys. Rev. B: Condens. Matter*, 1998, **57**, 14667.
32 B. Paulus, K. Rosciszewski, N. Gaston, P. Schwerdtfeger and H. Stoll, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2004, **70**, 165106.
33 B. Paulus, *Phys. Rep.*, 2006, **428**, 1.
34 C. Müller and D. Usvyat, Accurate calculations of the cohesive energy in molecular crystals: the role of embedding in the convergence of the incremental expansion, *J. Chem. Phys.*, 2011, submitted.
35 C. Müller, D. Usvyat and H. Stoll, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**, 245136.
36 M. Marsman, A. Grüneis, J. Paier and G. Kresse, *J. Chem. Phys.*, 2009, **130**, 184103.
37 A. Grüneis, G. Booth, M. Marsman, J. Spencer, A. Alavi and G. Kresse, *J. Chem. Theory Comput.*, 2011, **7**, 2780.
38 L. Maschio, D. Usvyat, F. R. Manby, S. Casassa, C. Pisani and M. Schütz, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **76**, 075101.

39 L. Maschio and D. Usvyat, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **78**, 073102.

40 L. Maschio, *J. Chem. Theory Comput.*, 2011, **7**, 2818.

41 D. Usvyat and M. Schütz, *J. Phys.: Conf. Ser.*, 2008, **117**, 012027.

42 R. F. W. Bader, *Atoms in Molecules—A Quantum Theory*, Oxford University Press, Oxford, UK, 1990.

43 C. Gatti, V. R. Saunders and C. Roetti, *J. Chem. Phys.*, 1994, **101**, 10686.

44 M. Cooper, *Adv. Phys.*, 1971, **20**, 643.

45 J. R. Hart and A. J. Thakkar, *Int. J. Quantum Chem.*, 2005, **102**, 673.

46 R. R. Zope, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2000, **62**, 064501.

47 A. Erba, M. Itou, Y. Sakurai, R. Yamaki, M. Ito, S. Casassa, L. Maschio, A. Terentjevs and C. Pisani, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**, 125208.

48 C. Pisani, M. Itou, Y. Sakurai, R. Yamaki, M. Ito, A. Erba and L. Maschio, *Phys. Chem. Chem. Phys.*, 2011, **13**, 933.

49 C. Pisani, A. Erba, S. Casassa, M. Itou and Y. Sakurai, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **84**, 245102.

50 M. Lorenz, D. Usvyat and M. Schütz, *J. Chem. Phys.*, 2011, **134**, 094101.

51 See ESI† for the parameters of the calculations.

52 M. Lorenz, D. Usvyat, L. Maschio and M. Schütz, Local *ab initio* methods for calculation optical band gaps in periodic systems: II. Periodic density fitted Configuration Interaction Singles method for solids, *J. Chem. Phys.*, 2011, to be submitted.

53 D. Usvyat, L. Maschio, C. Pisani and M. Schütz, *Z. Phys. Chem.*, 2010, **224**, 441.

54 J. W. Boughton and P. Pulay, *J. Comput. Chem.*, 1993, **736**, 14.

55 R. Orlando, R. Dovesi, C. Roetti and V. Saunders, *Chem. Phys. Lett.*, 1994, **228**, 225.

56 C. Müller and K. Hermansson, *Surf. Sci.*, 2009, **603**, 3329.

57 D. Usvyat, K. Sadeghian, L. Maschio and M. Schütz, Geometrical frustration of argon monolayer on MgO (110) surface: An accurate periodic *ab initio* study, *Phys. Rev. B: Condens. Matter*, 2011, to be submitted.

58 E. Voloshina, D. Usvyat, M. Schütz, Y. Dedkov and B. Paulus, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12041.

59 R. Martinez-Casado, G. Mallia, D. Usvyat, L. Maschio, S. Casassa, M. Schütz and N. M. Harrison, *J. Chem. Phys.*, 2011, **134**, 014706.

60 R. Martinez-Casado, G. Mallia, D. Usvyat, L. Maschio, S. Casassa, M. Schütz and N. M. Harrison, *Phys. Chem. Chem. Phys.*, 2011, **13**, 14750.

61 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.

62 M. Schütz, G. Rauhut and H.-J. Werner, *J. Phys. Chem. A*, 1998, **102**, 5997.

63 D. Usvyat, Van der Waals dispersion in the local correlation energy partitioning, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2011, submitted.

64 R. Martinez-Casado, D. Usvyat, G. Mallia, L. Maschio, S. Casassa, M. Schütz, N. M. Harrison, Scattering of helium on MgO(100)-surface: diffraction peak intensities and bound states from first principles, 2011, in preparation.

65 J. Langlet, J. Caillet, J. Bergès and P. Reinhardt, *J. Chem. Phys.*, 2003, **118**, 6157.

66 A. Heßelmann, *J. Chem. Phys.*, 2008, **128**, 144112.

67 A. Tkatchenko, R. A. DiStasio, M. Head-Gordon and M. Scheffler, *J. Chem. Phys.*, 2009, **131**, 094106.

68 E. Zaremba and W. Kohn, *Phys. Rev. B: Solid State*, 1976, **13**, 2270.

69 A. K. Geim and K. S. Novoselov, *Nat. Mater.*, 2007, **6**, 183.

70 R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796.

71 R. Demichelis, B. Civalleri, P. D'Arco and R. Dovesi, *Int. J. Quantum Chem.*, 2010, **110**, 406.

72 G. L. F. Bergaya and B. K. G. Theng, *Handbook of Clay Science*, Elsevier, Amsterdam, 2006.

73 W. Olbricht, N. D. Chatterjee and K. Miller, *Phys. Chem. Miner.*, 1994, **21**, 36.

74 R. A. Robie and B. S. Hemingway, *Am. Mineral.*, 1984, **69**, 298.

75 A. Hermann and P. Schwerdtfeger, *Phys. Rev. Lett.*, 2008, **101**, 183005.

76 R. Podeszwa, B. M. Rice and K. Szalewicz, *Phys. Rev. Lett.*, 2008, **101**, 115503.

77 O. Sode, M. Keçili, S. Hirata and K. Yagi, *Int. J. Quantum Chem.*, 2009, **109**, 1928.

78 S. J. Nolan, P. J. Bygrave, N. L. Allan and F. R. Manby, *J. Phys.: Condens. Matter*, 2010, **22**, 074201.

79 S. Wen and G. J. O. Beran, *J. Chem. Theory Comput.*, 2011, **7**, 3733.

80 A. Erba, L. Maschio, S. Salustro and S. Casassa, *J. Chem. Phys.*, 2011, **134**, 074502.

81 A. Erba, L. Maschio, C. Pisani and S. Casassa, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **84**, 012101.

82 L. Maschio, D. Usvyat and B. Civalleri, *CrystEngComm*, 2010, **12**, 2429.

83 L. Maschio, B. Civalleri, P. Ugliengo and A. Gavezzotti, *J. Phys. Chem. A*, 2011, **115**, 11179.

84 A. Erba, S. Casassa, L. Maschio and C. Pisani, *J. Phys. Chem. B*, 2009, **113**, 2347.

85 J. I. Lunine and D. J. Stevenson, *Icarus*, 1987, **70**, 61.

86 J. S. Loveday, R. J. Nelmes, M. Guthrie, S. A. Belmonte, D. R. Allan, D. D. Klug, J. S. Tse and Y. P. Handa, *Nature*, 2001, **410**, 661.

87 L. Maschio, D. Usvyat, M. Schütz and B. Civalleri, *J. Chem. Phys.*, 2010, **132**, 134706.

88 F. Bernardi and S. F. Boys, *Mol. Phys.*, 1970, **19**, 553.

89 C. Pisani, R. Dovesi, A. Erba and P. Giannozzi, in *Modern Charge Density Analysis*, ed. C. Gatti and P. Macchi, Springer, Berlin, 2011.

90 T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic Structure Theory*, John Wiley and Sons, Chichester, UK, 2000.

91 T. S. Koritsanszky and P. Coppens, *Chem. Rev.*, 2001, **101**, 1583.

92 P. Pattison, W. Weyrich and B. G. Williams, *Solid State Commun.*, 1977, **21**, 967.

93 A. J. Thakkar, A. M. Simas and V. H. Smith Jr., *Chem. Phys.*, 1981, **63**, 175.

94 P. J. Becker, J. M. Gillet, P. Cortona and S. Ragot, *Theor. Chem. Acc.*, 2001, **105**, 284.

95 Y. Tanaka, Y. Sakurai, A. T. Stewart, N. Shiotani, P. E. Mijnarends, S. Kaprzyk and A. Bansil, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2001, **63**, 045120.

96 A. Erba, C. Pisani, S. Casassa, L. Maschio, M. Schütz and D. Usvyat, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **81**, 165108.

97 A. Erba and C. Pisani, *J. Comput. Chem.*, 2012, DOI: 10.1002/jcc.22907.

98 R. S. Mulliken, *J. Chem. Phys.*, 1955, **23**, 1841.

99 The values of Mulliken charges obtained with a local density approximation (namely, LDA[107]) to the DFT are $+1.52$ and -0.76; with the popular hybrid scheme B3LYP[108] are $+1.82$ and -0.91 for the silicon and oxygen atoms, respectively.

100 E. R. Davidson, *J. Comput. Phys.*, 1975, **17**, 87.

101 D. Usvyat, B. Civalleri, L. Maschio, R. Dovesi, C. Pisani and M. Schütz, *J. Chem. Phys.*, 2011, **134**, 214105.

102 A. Lichanot, M. Gelize, C. Larrieu and C. Pisani, *J. Phys. Chem. Solids*, 1991, **52**, 1155.

103 L. Valenzano, Y. Noel, R. Orlando, C. Zicovich-Wilson, M. Ferrero and R. Dovesi, *Theor. Chem. Acc.*, 2007, **117**, 991.

104 S. Hirata, M. Head-Gordon and R. J. Bartlett, *J. Chem. Phys.*, 1999, **111**, 10774.

105 A. Izmaylov and G. Scuseria, *J. Chem. Phys.*, 2008, **129**, 034101.

106 S. Schmitt-Rink, D. S. Chemla and D. A. B. Miller, *Phys. Rev. B*, 1985, **32**, 6601.

107 J. P. Perdew and A. Zunger, *Phys. Rev. B*, 1981, **23**, 5048.

108 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.