Fragment-Based Restricted Active Space Configuration Interaction with Second-order corrections Embedded in Periodic Hartree-Fock Wavefunction

Hung-Hsuan Lin,[†] Lorenzo Maschio,[‡] Daniel Kats,[¶] Denis Usvyat,[*,§] and Thomas Heine[*,†]

[†]*Theoretische Chemie, Technische Universität Dresden, Germany*

[‡]*Dipartimento di Chimica, Università di Torino, Torino, Italy*

[¶]*Max-Planck-Institut für Festkörperforschung, Stuttgart, Germany*

[§]*Institut für Chemie, Humboldt-Universität zu Berlin, Germany*

E-mail: denis.usvyat@hu-berlin.de; thomas.heine@tu-dresden.de

Phone: +49 (351) 463 37637. Fax: +49 (351) 463 35953

Abstract

We present a computational scheme for restricted-active-space configuration interaction (RASCI) calculations combined with second-order perturbation theory (RASCI-PT2) on a fragment of a periodic system embedded in the periodic Hartree-Fock (HF) wavefunction. This method allows one to calculate the electronic structure of localized strongly-correlated features in crystals and surfaces. The scheme was implemented via an interface between the *Cryscor* and *Q-Chem* codes. To evaluate the performance of the embedding method, we explored dissociation of fluorine atom from a lithium fluoride surface and partially fluorinated graphane layer. The results show that RASCI

and RASCI-PT2 embedded in periodic HF are able to produce well-behaved potential energy surfaces and accurate dissociation energies.

Introduction

Theoretical modeling plays an important role in the development of new materials. For solid state calculations, Kohn-Sham (KS) density functional theory (DFT) is presently the most popular electronic structure method due to a good balance between accuracy and computational cost. However, standard KS-DFT within currently available functional approximations is known to fail for multireference systems[1,2] e.g., transition metal complexes, open-shell defects, systems far from equilibrium, etc. In this situation, good accuracy can be reached with the so-called multireference methods, which are, however, computationally very expensive and hence applications to periodic systems are limited.

In solids, multireference features have often local character, e.g. a local excitation, a vacancy, a catalytic reaction on a surface, etc. One possible approach in this case is to build a finite cluster by carving out the region of interest from the bulk. In this way one can employ molecular methods implemented in existing molecular quantum-chemical packages. Unfortunately, validity of the finite cluster model strongly depends on the type of the system. It is in principle applicable to either purely covalent solids, where the dangling bonds can be terminated with hydrogen atoms, or purely ionic ones, where no bonds need to be cut. However, for systems with a complex binding pattern or strongly disperse frontier bands the cluster approach is often unreliable. Besides, the long-range electrostatic interactions, present in solids, are missing in isolated (non-embedded) clusters, which can have a profound effect on the results.

Embedding approaches can be seen as an extension of the cluster model to study local effects in crystals. Several embedding methods have been developed.[3–14] They can be categorized into four classes: wavefunction partitioning, density-based, high-level-cluster

embedded in low-level-periodic model and quantum/classical force field approaches. These embedding classes can also be combined in hybrid schemes. As an example of density-based embedding one can mention the WFT-in-DFT embedding models, pioneered by Carter and co-workers.[6,15,16] Since then different flavors of DFT embedding have been developed and applied to molecular[9,11,17–26] and periodic systems.[15,16,27,28] Wavefunction partitioning and the related WFT-in-WFT embedding scheme can be seen as an alternative to the DFT embedding. This approach is presently a very active field of research.[14,17,29–34] One of those methods is fragment-based direct-local-ring coupled-cluster doubles (d-LrCCD) embedded in the periodic Hartree-Fock wavefunction, developed by some of us.[35] Although d-LrCCD itself does not guarantee more accurate results than MP2, it is much more stable for crystals with small band gap. For such systems it becomes instrumental to build hierarchical schemes that include higher-order corrections, evaluated on finite clusters (i.e., a hybrid embedding scheme, see above).[36,37] Very recently the periodic HF embedding was employed in conjunction with the Full CI quantum Monte Carlo (FCIQMC) technique.[38]

In this work we continue exploring the fragment approach with the periodic HF embedding. We present an interface for the second order perturbation theory based on the restricted-active-space configuration interaction (RASCI): RASCI-PT2 embedded in the periodic Hartree-Fock wavefunction. There are several flavors of the RASCI method; in this work we employ RASCI with one-hole and one-particle approximation. This variant of RASCI is an approximation to CASSCF (within the same active space RAS2), as the orbital relaxation is not performed in RASCI and is only partially taken into account by the single excitations from RAS1 and to RAS3. However, due to its computational simplicity RASCI allows for faster calculations and somewhat bigger active spaces than CASSCF.

The purpose of this paper is to present the embedded RASCI-PT2 approach and to assess its performance on a particular example of F atom dissociation from a (001) lithium fluoride (LiF) surface and from a partially fluorinated graphane (HF-graphane). In these systems the dynamic correlation is substantial due to the presence of the fluorine atom, and the amount

of the static correlation can be tuned by increasing the C−F or Li−F bond distance up to dissociation. In the HF-graphane case, RASCI-PT2 is compared against the distinguishable cluster[39] (DC) approximated coupled cluster with singles doubles and triples (DC-CCSDT) method,[40] which is expected to deliver accurate results for fluorine dissociation.

Theoretical Background

Embedding Workflow

The embedding scheme used here is related to that of Masur and coworkers,[35] where it was used in conjunction with local ring-CCD. In contrast to that method, here we use local orbitals for the fragment definition only, and once the fragment is defined the orbitals are canonicalized within the fragment. The detailed specification of the current model is to be presented elsewhere[38] and will be not reiterated here. We provide only the general concept, which can be summarized as follows:

- The orbital space of the fragment is defined by a group of localized occupied and virtual orbitals that are spatially close to the site of interest. This allows for a seamless specification of the fragment and its embedding into the HF mean-field from the rest of the crystal. The occupied orbitals – Wannier functions (WFs) – are mutually orthogonal and correspond to the periodic HF solution. The virtual space is spanned by the so called projected atomic orbitals (PAOs), i.e. AOs projected from the occupied space of the periodic HF solution. In this way the orbital space of the fragment is by construction orthogonal to the HF occupied space of the embedding environment.

- Once the occupied and virtual spaces of the fragment are defined, the Fock matrix block, corresponding to the fragment orbitals, is cut out from the periodic Fock matrix and diagonalized. The resulting eigenfunctions are the canonical occupied and virtual orbitals of the fragment. These canonical orbitals are then used as the basis for the

in-fragment post HF treatment. The active spaces for the RASCI and RASCI-PT2 procedures are chosen on the basis of the eigenvalues of these Fock matrices (i.e. orbital energies).

- The matrix elements of the fragment's Hamiltonian in the basis of the fragment's canonical orbitals are calculated using the periodic *Cryscor* code, where the AO support extends into the environment region. The procedure consists of
  - evaluation of the electron repulsion integrals by means of the periodic local density fitting,[41,42]
  - construction of the fragment's core Hamiltonian h^{frag} from the periodic Fock matrix by subtracting the Coulomb and exchange contributions belonging to the fragment.

  This procedure thus implicitly includes the Coulomb and exchange field of the environment in h^{frag}.

- The formal nuclear repulsion energy of the fragment $E_{\text{nuc}}^{\text{frag}}$ is defined such that the fragment's HF energy, computed using h^{frag} and electron-electron repulsion integrals in the basis of the fragments occupied orbitals, coincides with the periodic HF energy. This choice of $E_{\text{nuc}}^{\text{frag}}$ implies that the whole fragment energy (HF energy plus the correlation correction) can also be interpreted as the energy per cell, where the fragment approximation restricts the correlated orbital space to the fragment orbitals and neglects the correlated couplings beyond the fragment.

The workflow is outlined in Figure 1. The implementation of this scheme involves three quantum chemistry packages, i.e., *Crystal17*,[43] *Cryscor*,[44] and *Q-Chem*[45] The final goal is to run the RASCI-PT2 method using a set of orbitals from the converged periodic HF calculation as a basis set, and the Hamiltonian that includes the periodic mean field.

In the initial stage the periodic restricted closed-shell HF wavefunction is calculated

by *Crystal17*. WFs, which are then constructed by localization of converged HF Bloch functions,[46] as well as the Fock matrix and the relevant auxiliary quantities, like the basis set, geometry, etc, are passed to *Cryscor*.[44] The localized virtual orbitals, i.e. PAOs, are computed on the *Cryscor* side,[47] as well as other quantities such as the electron repulsion integrals (ERIs), the fragment's core Hamiltonian h^{frag} and $E^{\text{frag}}_{\text{nuc}}$. The ERIs, h^{frag} and $E^{\text{frag}}_{\text{nuc}}$ together with the number of electrons (for neutral fragments: the half of number of the fragment's WFs) form the essential data needed for the fragment RASCI-PT2 treatment. They are passed to the molecular code *Q-Chem*, which proceeds with the RASCI and RASCI-PT2 calculation.

Technically, *Q-Chem* receives the one- h^{frag}_{pq} and two-electron part $(pq|rs)$ of the Hamiltonian in the basis of the fragments orbitals p, q, r and s in the format equivalent to that of Ref. 48. These orbitals take the role of the basis functions for the fragment, and h^{frag}_{pq} and $(pq|rs)$ substitute the corresponding AO-based quantities of *Q-Chem*. Since all the transformations from the AO basis and $E^{\text{frag}}_{\text{nuc}}$ have already been evaluated on the periodic side, the atoms and the corresponding basis functions themselves are irrelevant for the *Q-Chem* part. The basis-orbital overlap matrix is replaced with an identity matrix as the fragment canonical orbitals are orthonormal. *Q-Chem* starts with a HF calculation, which, in case of a singlet state, converges in one iteration, as the fragment's orbitals by construction obey the Brillouin condition for the fragment's Fock matrix. At this point, RASCI and the subsequent second-order perturbative correction are ready to go. The RASCI and RASCI-PT2 approaches of the *Q-Chem* program can be found in Refs. 49–54. We specify it in some detail in the next subsection.

RASCI and RASCI-PT2 with hole and particle approximation

Generally in the RASCI method,[55] the active orbitals of fragment are divided into three subspaces, i.e. RAS1, RAS2, and RAS3. The RAS2 subspace is similar to the complete active space of the CASSCF or CASCI method and allows all possible configurations of its

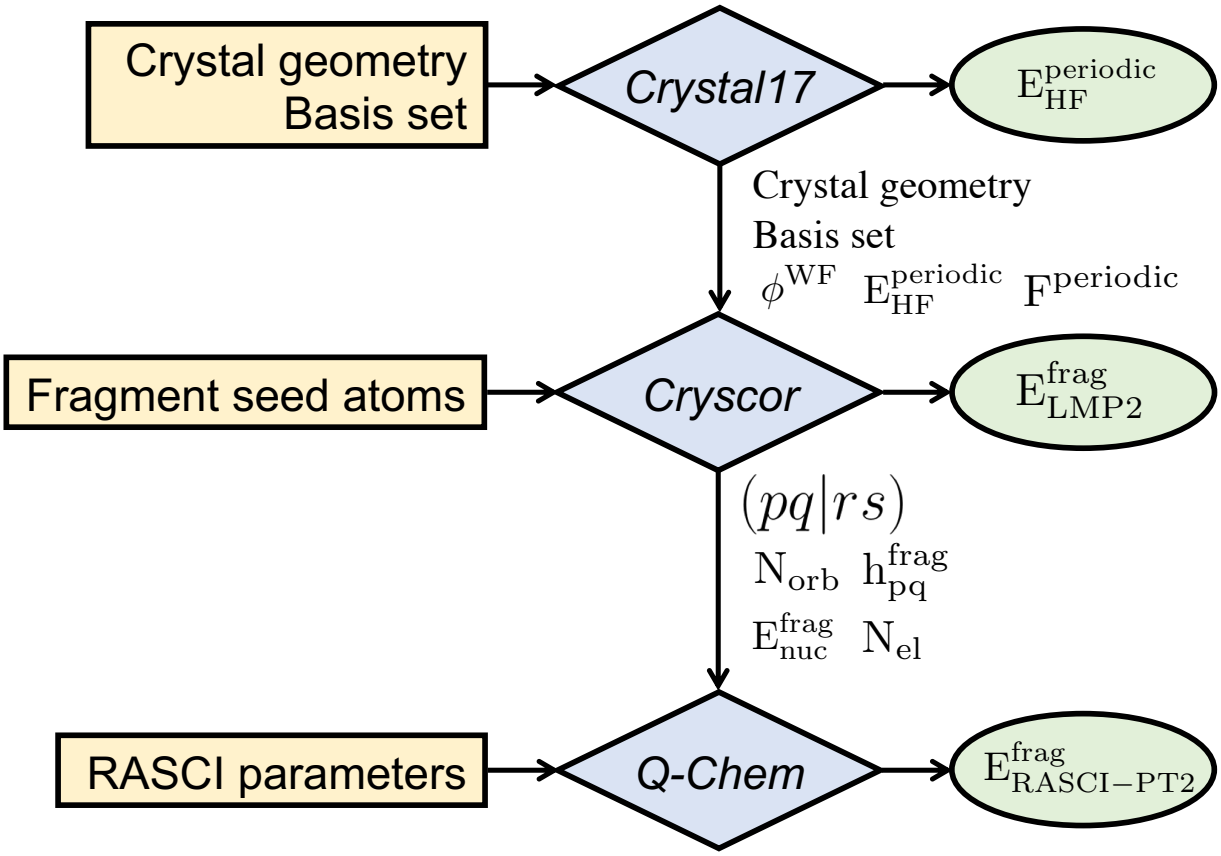

Figure 1: The workflow for the embedded-fragment calculations. $E_{\rm HF}^{\rm periodic}$ is the periodic HF energy; $F^{\rm periodic}$ is the periodic Fock matrix (in the AO basis); ϕ^{WF} are the occupied WFs; $(pq|rs)$ are the two-electron integrals in the basis of fragment canonical orbitals denoted by the indices p, q, r and s; $h_{pq}^{\rm frag}$ is the matrix of the fragment's one-electron operator, that includes the HF mean field from the embedding; $E_{\rm nuc}^{\rm frag}$ is the formal nuclear energy to equalize the fragment and periodic HF energies; $N_{\rm el}$ – number of electrons in the fragment; $N_{\rm orb}$ – number of orbitals in the fragment; $E_{\rm RASCI-PT2}^{\rm frag}$ is the target RASCI-PT2 energy of the embedded fragment.

n electrons in m orbitals. In contrast, in the RAS1 and RAS3 subspaces the occupations are restricted to a maximum number of holes and electrons, respectively. Following the formalism of Ref. 54 we write the excitation operator as

$$\hat{R} = \hat{r}^0 + \hat{r}^h + \hat{r}^p + \hat{r}^{hp} + \hat{r}^{2h} + \hat{r}^{2p} + ..., \tag{1}$$

where $\hat{r}^0$ generates the full determinantal space within RAS2, while further operators extend this space by certain classes of determinants. The superscripts h and p represent, respectively, hole and electron excitations, such that, for example, $\hat{r}^h$ generate configurations with one hole in RAS1, $\hat{r}^p$ - with one electron in RAS3, etc.

In this work we employ the hole and particle approximation within RASCI, which truncates the operator (1) to the first three terms, or in other words allows only one hole or one electron in RAS1 or RAS3, respectively. Since there are only single excitations from RAS1 or to RAS3, the size extensivity issues are not as severe as in RASCI (or RASSCF) with two holes and electrons in RAS1 and RAS3 spaces or as in MRCI. Furthermore, with the hole and particle approximation it usually becomes unproblematic to extend RAS1 and RAS3 to the full occupied and virtual spaces, respectively. In this form, RASCI can be seen as an approximation to CASSCF, where the single excitations to some extent mimic orbital relaxations.

The RASCI method with the hole and particle approximation is not able to capture most of the dynamical correlation. With a standard active space RASCI only recovers less than 20 % of the CCSD(T) correlation energy.[54] Therefore for quantitative accuracy RASCI is to be followed by a perturbative treatment. We employ the state-specific multi-reference perturbation theory of Ref. 54 as implemented in the *Q-Chem* code.

The method employs the following zeroth order Hamiltonian:

$$\hat{H}_0 = E_0 \left|0\right\rangle\left\langle 0\right| + \sum_k E_k \left|k\right\rangle\left\langle k\right|, \tag{2}$$

where $|0\rangle$ and E_0 are, respectively, the Hamiltonian eigenfunction and eigenvalue for the state of interest in the space generated by $\hat{r}^0$. In other words, $|0\rangle$ and E_0 correspond to the CASCI solution in the RAS2 space. The configurations $|k\rangle$ form the orthogonal complement to $|0\rangle$ and (together with $|0\rangle$) span the space generated by the operators of eq. (1) that generate at most two holes in RAS1 and two electrons in RAS3. In other words the RASCI-PT2 wavefunction includes up to the doubly excited configurations with respect to the determinants of RAS2. For the energies E_k we use the generalized Davidson-Kapuy partitioning,[56] which represents them as differences of the diagonal elements of a generalized Fock matrix, see equations (6) and (8) of Ref. 54.

Since the zeroth-order Hamiltonian corresponds to the CASCI solution in RAS2 rather than the actual RASCI solution, the determinants with a hole in RAS1 or a particle in RAS3 occur among the $|k\rangle$ states. In order to avoid double counting, the contributions from these configurations must be excluded from the second order energy correction, which then takes the form:

$$E^{RASCI-PT2} = E^{RASCI}_{0\oplus H\oplus P} - \sum_{k\neq H,P} \frac{\left|\left\langle k \left| \hat{H} \right| 0 \right\rangle\right|^2}{E_k - E_0}, \tag{3}$$

where H and P denote these hole and particle spaces, respectively.

Results and discussion

LiF

As a first test we consider dissociation of a fluorine atom from a surface of an ionic LiF crystal. Dissociation of a LiF molecule is a standard test for a multireference treatment to study the ionic-neutral crossing. To test our embedding method we consider the periodic analog of this system. To model the (001) surface of LiF we used a three-layer slab with a supercell containing 8 LiF units per layer. The model is shown in Figure 2. The structure was relaxed at B3LYP/DZVP level and is explicitly given in Supporting Information (see Figure

S2). For the embedded RASCI and RASCI-PT2 calculations we considered a fragment with one fluorine atom surrounded by five lithium atoms, as also shown in Figure 2.

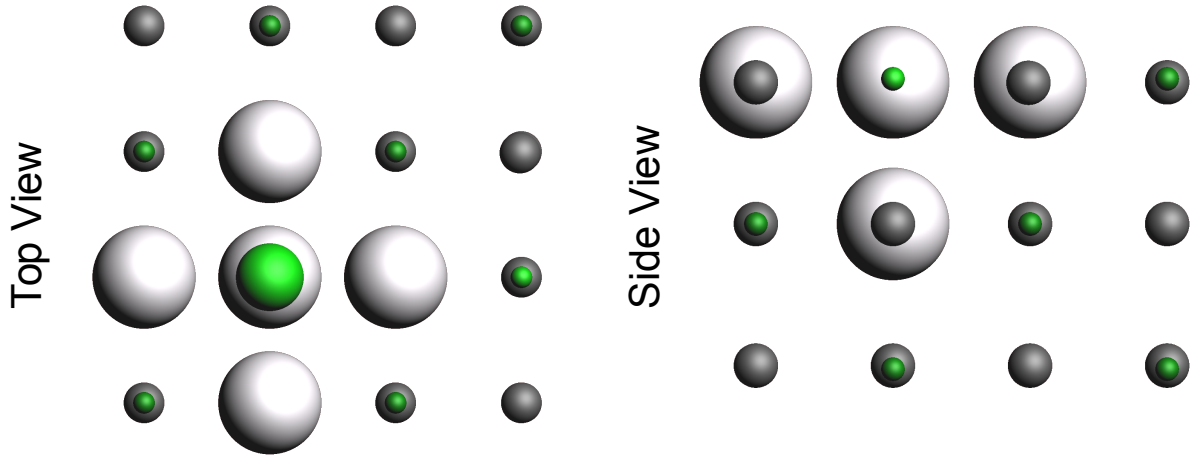

Figure 2: The supercell of the three-layer lithium-fluoride slab and the fragment used in this study. The grey and green spheres refer to lithium and fluorine atoms, respectively. The fragment is marked by larger spheres.

The ground and excited state potential energy curves are presented in Figure 3. As expected, Hartree-Fock remains in the ionic state throughout dissociation. At the same time, both RASCI and RASCI-PT2 correctly reproduce the transition of the ground state from the ionic state via an avoided crossing to the triply degenerate neutral state. This suggests that the embedded fragment approach provides at least a qualitatively correct description of dissociation in this system.

Interestingly, from Figure 3 one can conclude that the dynamic correlation has only a weak effect on the position of the avoided crossing. According to the results of Ref. 57, in the LiF molecule the CASSCF prediction for the crossing can noticeably deviate from that of CASPT2. However, the CASSCF results depend strongly on the choice of the active space, and for an active space comparable to that used here, the differences between CASSCF and CASPT2 crossing positions are also quite small.[57] We finally note, that compared to RASCI, RASCI-PT2 predicts lower excitation energies at dissociation, but higher excitation energies near the minimum.

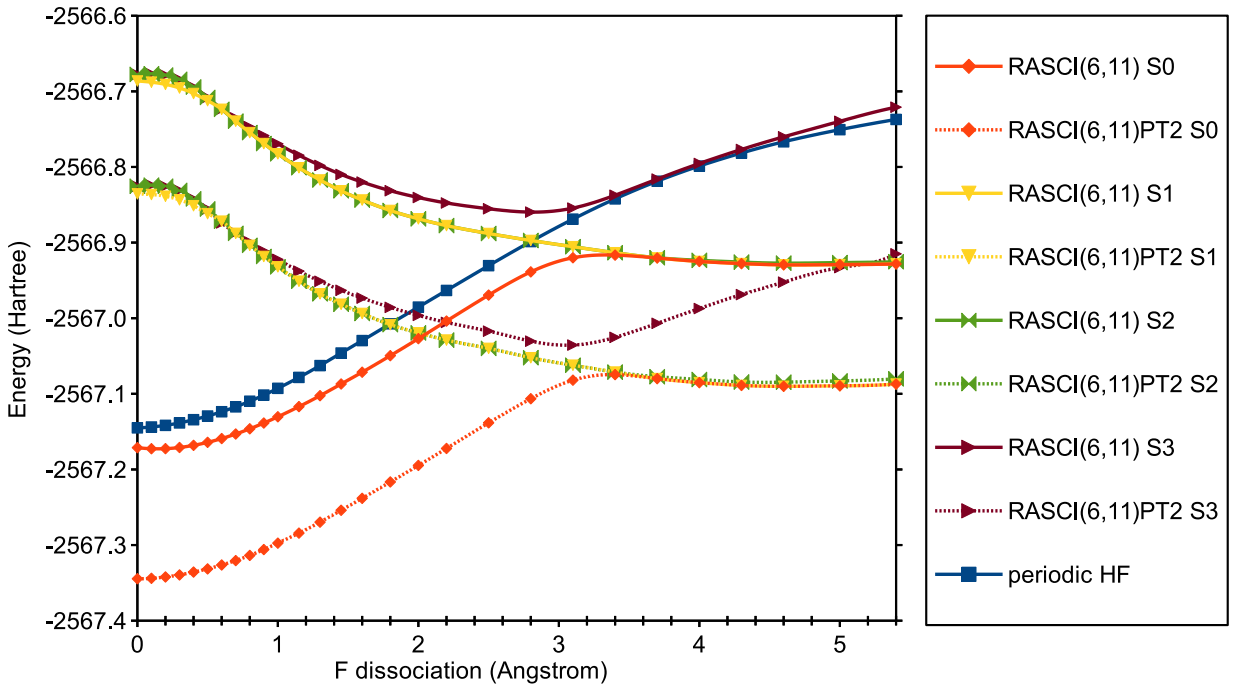

Figure 3: Potential energy curves for F dissociation from the (001) LiF surface, calculated with periodic HF and embedded-fragment RASCI and RASCI-PT2 methods. S0 denotes the ground state, while S1, S2, and S3 the corresponding excited states.

HF-graphane

As a second example we consider covalent-bond breaking. A single bond dissociation is a standard test to assess an ability of a method to accurately describe a multireference case. When a simple (e.g. hydrogen-hydrogen) covalent bond is stretched, the orbital energies of the occupied bonding HF orbital σ and the virtual anti-bonding HF orbital (σ^*) are getting close. Therefore, in order to describe the ground (singlet) state of such a system even qualitatively correctly, at least two configurations are needed in the wavefunction expansion.

In this work we explore a complicated case: dissociation of the C-F bond in a partially fluorinated graphane. Firstly, this is a periodic system and the simple bonding-antibonding orbital picture is not directly applicable, as the canonical orbitals are delocalized Bloch waves. Secondly, the ground state of an isolated fluorine atom is triply degenerate, so at dissociation several orbitals, not just two, become close in the energies. Finally, since the number of electrons in the vicinity of the bond is large (fluorine alone has six lone-pair valence

electrons), the dynamic correlation becomes absolutely essential for quantitative accuracy.

Specification of the model parameters

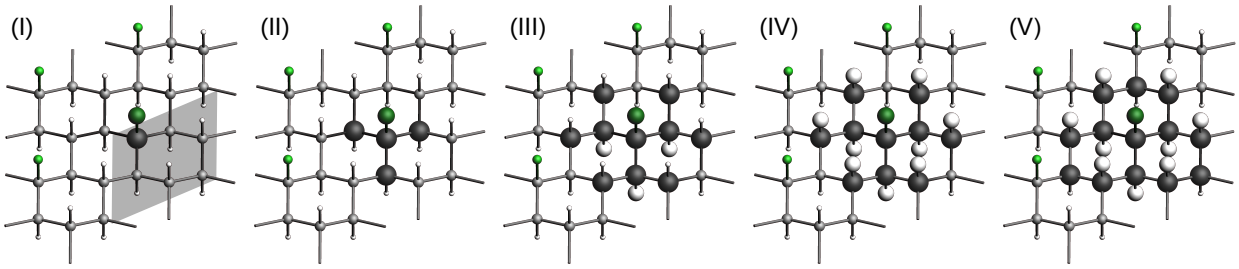

Figure 4: Partially fluorinated graphane and its fragments used in this study. The grey, white and green spheres refer to carbon, hydrogen, and fluorine atoms, respectively. The fragments are represented by the larger spheres. The shaded region highlights the unit cell.

The model is illustrated in Figure 4. The structure of periodic partially fluorinated graphane was optimized at the B3LYP/6-21G* level. The geometry can be found in the Supporting Information in the form of *Crystal17* input (see Figure S1). We considered several fragments with progressively increasing size: from 8 to 56 electrons. Figure 4 gives a graphical representation of the fragments by the nuclei denoted with larger spheres. The orbital space of each fragment includes PAOs, centered on these atoms, and WFs, which are either bonding or lone-pair WFs corresponding to these atoms. A detailed specification of the fragments is given in Table 1.

Table 1: Specification of the fragments. N_{el} is the number of electrons in the fragment (twice the number of WFs), N_{orb} is the number of orbitals in the fragment (both occupied and virtual). We also list the active spaces RAS2 that we employ for different fragments.

Fragment	Atoms	N_{el}	N_{orb}	RAS2	
				Type 1	Type 2
I	1 F, 1 C	8	30	(6,4)	(6,7)
II	1 F, 4 C	14	72	(10,6)	(10,9)
III	1 F, 10 C, 3 H	32	173	(22,12)	(10,12)
IV	1 F, 10 C, 9 H	44	207	(34,18)	
V	1 F, 13 C, 9 H	56	248	(42,22)	

The RASCI and RASCI-PT2 calculations were performed according to the above de-

scribed scheme that involves three codes: *Crystal17*, *Cryscor* and *Q-Chem*. The initial periodic HF calculation employed the 6-21G* basis set. Throughout the periodic stages of the calculations (HF, orbital localization, construction of PAOs) the 8x8 k-grids were used. For the fluorine dissociation, only the C-F distance was varied, the rest of the structure was kept frozen.

For the reference values we performed DC-CCSDT calculations for the first three fragments using an interface to the Molpro[58] and GeCCo[59] codes. For the smallest fragment, we also did MP2, CCSD and CCSDTQ, the first two with Molpro and the latter with the MRCC code,[60] interfaced with Molpro. In all these calculations the necessary quantities for high-level methods were also generated by *Cryscor*.

Similar to other multireference methods, RASCI is not black-box, as one needs to choose the active space. In our case this concerns RAS2, as RAS1 and RAS3 are then defined automatically to include the complete closed-shell and virtual spaces, respectively. RASCI is expected to be more sensitive to the choice of the active space than CASSCF, since it does not reoptimize the HF orbitals, which might be far from optimal in the multireference regime. In this work we test two types of active spaces, type 1: with plenty of electrons, but with just one empty orbital to excite to; and type 2: a standard active space with a moderate number of electrons in approximately the same number of orbitals. The exact specification of the active spaces for each fragment is also given in Table 1.

Potential energy curves

We start with the smallest fragment I, consisting of just two atoms, F and C, corresponding to the dissociating bond. The potential energy curves for this fragment are presented in Figure 5. As expected the standard single reference methods (here HF, MP2, CCSD) do not dissociate the fluorine atom even qualitatively correctly. The coupled cluster method with up to quadruple excitations (CCSDTQ) is able to dissociate single bonds and, moreover, accurately captures the dynamic correlation. It could be used as a reference, but it is very

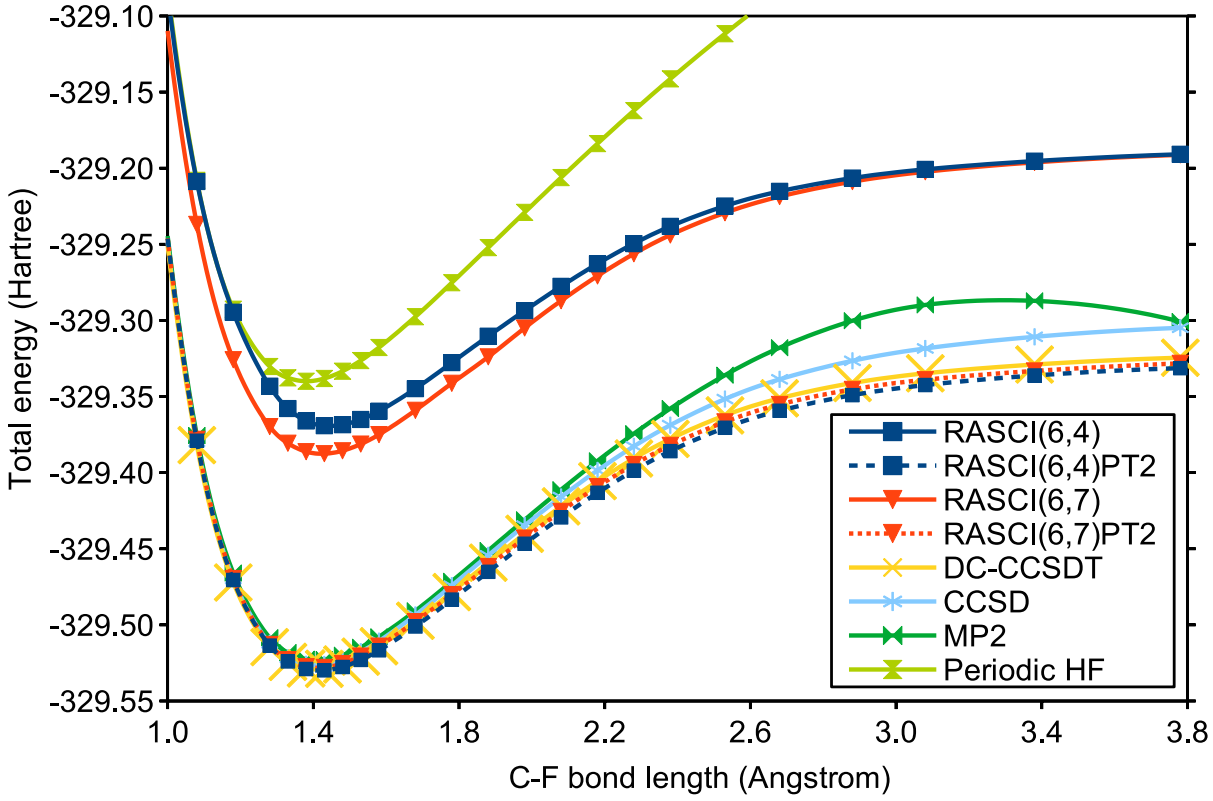

Figure 5: Potential energy curves for C-F bond dissociation for the fragment I, calculated with periodic HF and embedded-fragment MP2, CCSD, DC-CCSDT, RASCI, and RASCI-PT2 methods.

computationally expensive and is feasible only for the smallest fragment. Distinguishable cluster method with up to triples excitations DC-CCSDT is known to perform well for fluorine dissociation,[40] and indeed it deviates from CCSDTQ only in the sub-millihartree region along the whole curve (see supplementary information). We thus use DC-CCSDT as the reference for RASCI-PT2.

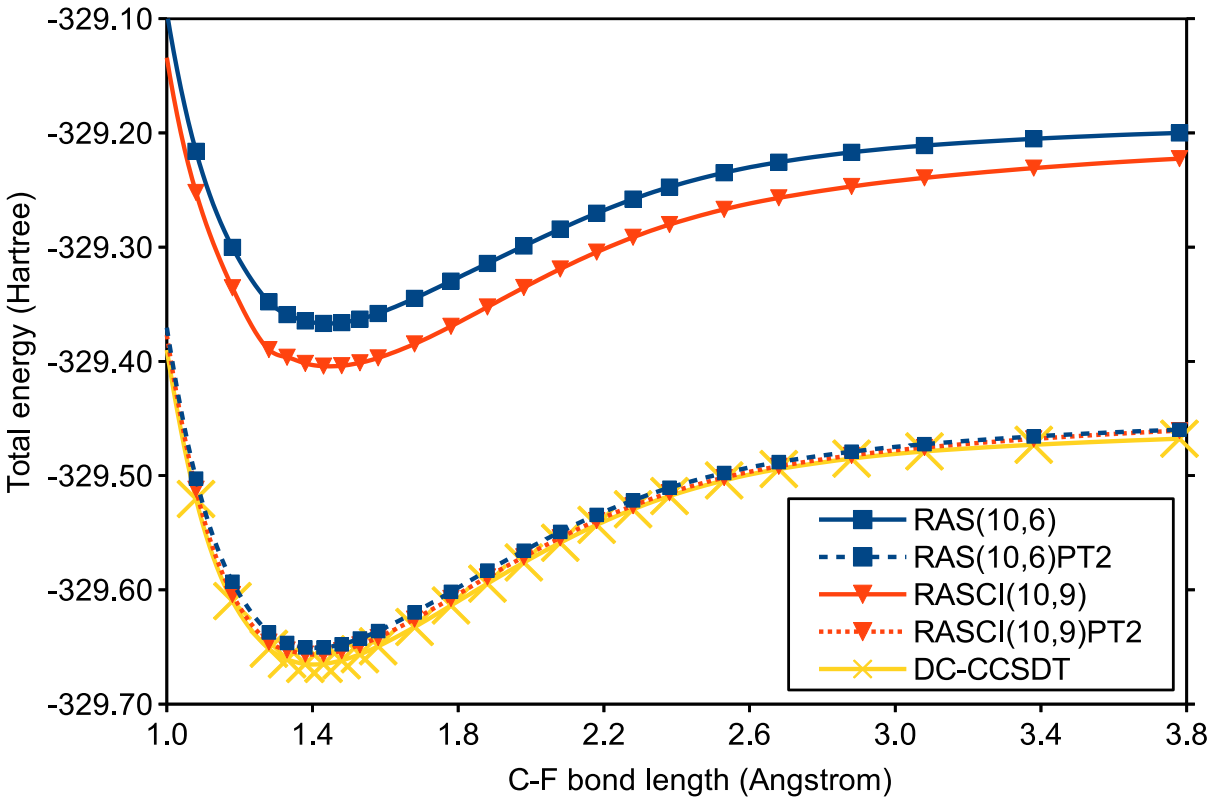

Figure 6: Potential energy curves for C-F bond dissociation for the fragment II, calculated with periodic HF and embedded-fragment RASCI, RASCI-PT2 and DC-CCSDT methods.

The multireference treatments RASCI and RASCI-PT2 provide a qualitatively correct dissociation for this fragment regardless of the active space choice. RASCI is clearly missing a big share of dynamic correlation, but RASCI-PT2 recovers it virtually completely for both active space choices (and even slightly overcorrelates). A similar picture is observed for the fragment II, presented in Figure 6. However, one can notice here that the RASCI-PT2 curves are already slightly above the DC-CCSDT one. In this fragment the deviations of the RASCI-PT2 total energies from those of DC-CCSDT, which are shown in Figure S3, are

around 2-20 millihartree, which is higher than in fragment I. One of the reasons for that is a general loss of size extensivity in a RASCI (or CASCI/CASSCF) energy if the active space is not expanded proportionally to the system size. We note, however, that in RASCI-PT2 this effect is not as pronounced as in RASCI itself. Besides, the the errors in the dissociation energies (Figure S4), which are much more informative than the total energies, are similar to those of fragment I.

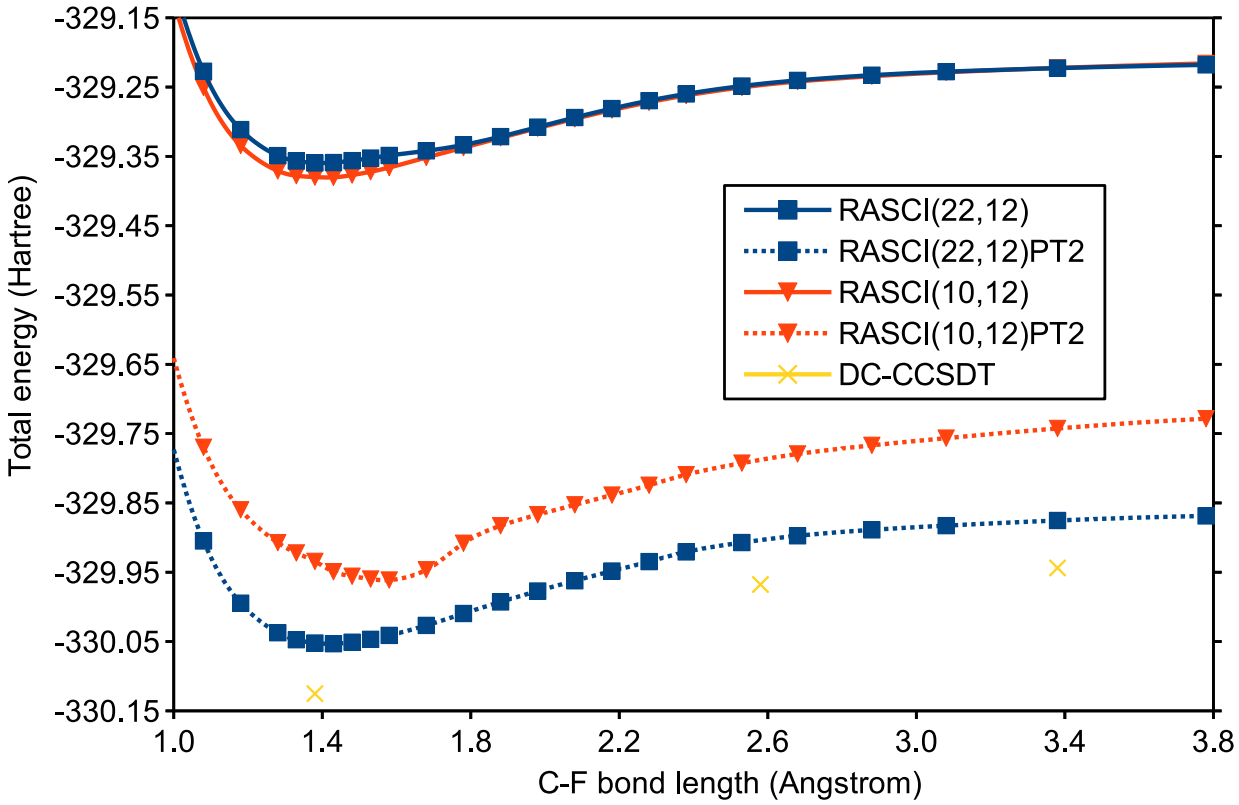

Figure 7: Potential energy curves for C-F bond dissociation for the fragment III, calculated with periodic HF and embedded-fragment RASCI, RASCI-PT2 and DC-CCSDT methods.

As is shown in Figure 7, for the fragment III the RASCI-PT2 potential energy curve, corresponding to the "standard" active space (10,12), exhibits some non-physical features. We were not able to get rid of these artifacts by varying the active space by an orbital or two. By visualizing the active orbitals (shown in Figures S7 and S8) we found both types of active space to be large enough to include all the relevant orbitals, i.e. sigma bonding and antibonding, p_x and p_y of the F atom. Therefore, the active space is adequate and absence

of orbital rotations in RASCI cannot be the origin of this issue.

To investigate this problem further, we considered a molecule C_4H_9F as finite cluster prototype of Fragment III. The molecule was cut out from the periodic structure with the dangling bonds terminated by hydrogen atoms. The results are shown in Figure S6. Similar to the fragment calculations, RASCI(10,12)PT2 manifests here non-physical artifacts. Switching to the spin-flip version of the RASCI-PT2 method,[61–63] which uses the triplet state as the reference, does not solve the problem. At the same time, both CASPT2(10,9) and CASCI(10,9)PT2 (the latter using the same RHF orbitals as in RASCI(10,9)PT2), calculated with Molpro, provide physically correct curves. This suggests that the problems in the RASCI(10,12)PT2 calculations are either caused by the RASCI-PT2 method itself or its implementation in Q-Chem.

Importantly, the active space of type 1 – (22,12) – does not cause such problems. Indeed, the RASCI(22,12)PT2 curve is smooth and parallel to the three available DC-CCSDT data points (for the fragment of this size DC-CCSDT in its present implementation is already quite expensive). As expected the deviations of the RASCI(22,12)PT2 total energies from DC-CCSDT grow with fragment size and are noticeably bigger than for fragment II (see Figure S3). But the error in the dissociation energies (Figure S4) remains in the millihartree region and is insensitive to the fragment size.

Convergence with fragment and supercell sizes

For chemistry, the total energies themselves are of limited interest, as any chemical process is governed by energy differences. In this study the relevant energy difference is the dissociation energy, which we define here as the difference between the total energies at $R_{\mathrm{C-F}} = 3.38$ Å and $R_{\mathrm{C-F}} = 1.38$ Å. The results for the dissociation energy computed on different fragments are given in Figure 8. The DC-CCSDT method is again used as a reference.

RASCI, despite the qualitatively correct description, quantitatively grossly underestimates the dissociation energy, especially so with the small active space (type 1). The

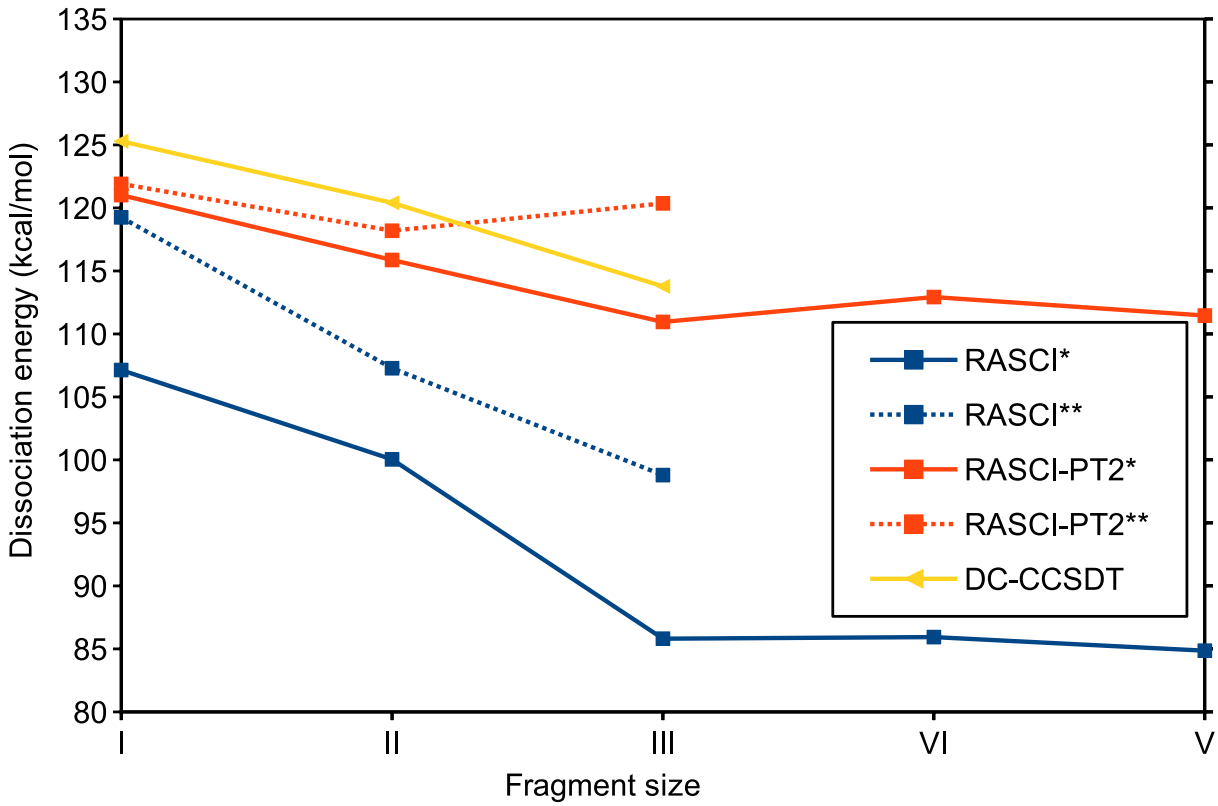

Figure 8: The RASCI, RASCI-PT2 and DC-CCSDT energies of the C-F bond dissociation. The active spaces of type 1 and 2 are marked with one and two stars, respectively. The illustration and specification of the fragment size can be found in Figure 4 and Table 1, respectively.

RASCI-PT2 dissociation energies are much more accurate. They agree with the reference values within a few kcal/mol, which is a standard uncertainty expected from a second-order perturbative treatment. Furthermore, the RASCI-PT2 results are less sensitive to the choice of the active space than RASCI, until it runs into the unphysical regime in fragment III with the active space of type 2. RASCI-PT2 with the active space of type 1, on the other hand, behaves robustly and can be used to explore the convergence with fragment size. Fragments IV and V are too big for DC-CCSDT in its present implementation, but RASCI-PT2 demonstrates that after fragment III the dissociation energy is essentially converged, at least within a few kcal/mol uncertainty.

Finally we investigate the dependence of the results on the supercell chosen. The supercell size is an important parameter of the model. Indeed a defect, in the case of dissociated fluorine atom, is periodically repeated with the period defined by the supercell. The fragment is defined to cover the defect in the reference cell, but the periodic replicas of the defect can still influence the fragment via the embedding. In our model the embedding remains at the uncorrelated HF level, which can be insufficient for an adequate description of the defect replicas.[64] One way to cope with this problem would be to increase the embedding level or to couple the correlated replicas of the fragments.[65,66] Yet, for extended periodic systems this is difficult. A much simpler procedure would be to "dilute" the defect in the crystal by increasing the supercell size until convergence is reached.

In this work in addition to the initial unit cell, we consider two supercells: 3x3 and 5x5 (both treated with a single Γ k-point in the Brillouin zone). The RASCI and RASCI-PT2 dissociation energies for fragments I-III and type 1 active space are summarized in the Figure 9 (and Table S6). There is dependence on the supercell size and with the choice of the initial unit cell, the influence of the replicas of the defect is non-negligible. However, for RASCI-PT2 the difference between 3x3 and 5x5 dissociation energies is already not greater than 2 kcal/mol, which can be considered converged. The convergence of RASCI is somewhat slower, but for the fragment III it also seems to be converged already with the 3x3 supercell.

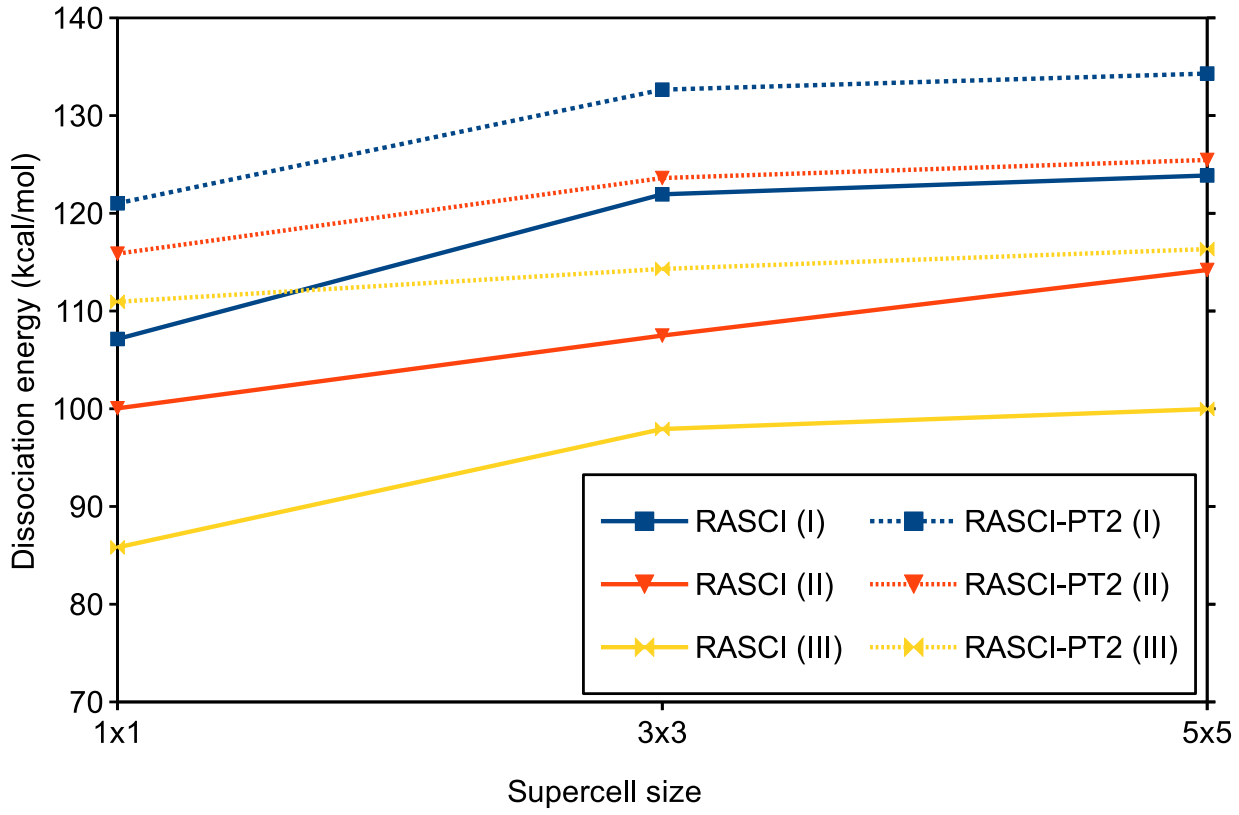

Figure 9: The RASCI and RASCI-PT2 dissociation energies for Fragments I, II, and III, 1x1, 3x3 and 5x5 supercells, and the active space of type 1.

Conclusions

In this work we presented an embedding scheme for investigation of local multireference features inside crystalline systems. The method allows for RASCI-PT2 treatment of fragments embedded in the periodic Hartree Fock wavefunction. Fragments are defined by groups of localized occupied and virtual orbitals, corresponding to the periodic Hartree-Fock solution. These orbitals serve as a basis set for the fragment-based multireference treatment. The embedding field, consisting of the Coulomb potential from the crystalline nuclei and the HF mean-field Coulomb and exchange potentials from the electrons outside the fragment, is effectively included in the one-electron Hamiltonian of the fragment. An embedded RASCI-PT2 treatment, based on RASCI with hole and particle approximation, captures both dynamic and static correlation in the region surrounding the site of interest and can deliver quantitatively accurate results at a relatively low for a multireference method computational cost. This makes it possible to apply this technique to rather extended fragments and probe the convergence of the results with the fragment size.

The method has been tested on a fluorine atom dissociation from a surface of LiF crystal and a partially fluorinated graphane. In the former case the qualitative picture of avoided crossing between the ionic and neutral state has been reproduced. For the latter system, a simple choice of the active space that includes most of valence electrons but just one virtual orbital to excite to is sufficient to reach a good accuracy for the dissociation energy with a deviation from DC-CCSDT of just a few kcal/mol. The convergence of the dissociation energy with fragment size was achieved at a 14-atom fragment (fragment III) and 3x3 supercell. To estimate the computational cost of this technique: a single point embedded RASCI-PT2 calculation for the fragment III with the (22,12) active space on a single thread Intel of Xeon E5-2640 v4 @ 2.40GHz took about 6.5 minutes. We believe that this method will be a reliable and computationally affordable tool to study multireference problems in solids.

Acknowledgement

We acknowledge support from Volkswagenstiftung and Deutsche Forschungsgemeinschaft (Grants HE 3543/31-1 and US 103/1-2). DK acknowledges financial support from the Max-Planck Society.

Supporting Information Available

Detailed specification of the computational parameters and the geometry as well as the values of the calculated energies can be found in Supplementary Information

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Science* **2008**, *321*, 792–794.

(2) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. *Chem. Rev.* **2011**, *112*, 289–320.

(3) Sauer, J. *Chem. Rev.* **1989**, *89*, 199–255.

(4) Wesolowski, T. A.; Warshel, A. *J. Phys. Chem.* **1993**, *97*, 8050–8053.

(5) Whitten, J. L.; Yang, H. *Surf. Sci. Rep.* **1996**, *24*, 55–57.

(6) Govind, N.; Wang, Y.; da Silva, A.; Carter, E. *Chem. Phys. Lett.* **1998**, *295*, 129–134.

(7) Trail, J. R.; Bird, D. M. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2000**, *62*, 16402–16411.

(8) Jug, K.; Bredow, T. *J. Chem. Theory Comput.* **2004**, *25*, 1551–1567.

(9) Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesolowski, T. A. *J. Chem. Phys.* **2005**, *122*, 094115.

(10) Jacob, C. R.; Neugebauer, J.; Visscher, L. *J. Comput. Chem.* **2008**, *29*, 1011–1018.

(11) Gomes, A. S. P.; Jacob, C. R.; Visscher, L. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5353–5362.

(12) Huang, P.; Carter, E. A. *Annu. Rev. Phys. Chem.* **2008**, *59*, 261–290.

(13) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. *Chem. Rev.* **2015**, *115*, 5678–5796.

(14) Paulus, B.; Rosciszewski, K. *Chem. Phys. Lett.* **2004**, *394*, 96–100.

(15) Klüner, T.; Govind, N.; Wang, Y. A.; Carter, E. A. *J. Chem. Phys* **2002**, *116*, 42.

(16) Huang, P.; Carter, E. A. *J. Chem. Phys.* **2006**, *125*, 084102.

(17) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.

(18) Hégely, B.; Nagy, P. R.; Ferenczy, G. G.; Kállay, M. *J. Chem. Phys.* **2016**, *145*.

(19) Barnes, T. A.; Goodpaster, J. D.; Manby, F. R.; Miller, T. F. *J. Chem. Phys.* **2013**, *139*.

(20) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F. *J. Chem. Phys.* **2014**, *140*.

(21) Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. *J. Phys. Chem. A* **2014**, *118*, 9182–9200.

(22) Fornace, M. E.; Lee, J.; Miyamoto, K.; Manby, F. R.; III, T. F. M. *J. Chem. Theory Comput.* **2015**, *11*, 568.

(23) Culpitt, T.; Brorsen, K. R.; Hammes-Schiffer, S. *J. Chem. Phys.* **2017**, *146*.

(24) Chulhai, D. V.; Goodpaster, J. D. *J. Chem. Theory Comput.* **2017**, *13*, 1503–1508.

(25) Tamukong, P. K.; Khait, Y. G.; Hoffmann, M. R. *J. Phys. Chem. A* **2016**, *121*, 256–264.

(26) Shimazaki, T.; Kitaura, K.; Fedorov, D. G.; Nakajima, T. *J. Chem. Phys.* **2017**, *146*.

(27) Libisch, F.; Marsman, M.; Burgdörfer, J.; Kresse, G. *J. Chem. Phys.* **2017**, *147*.

(28) Chulhai, D. V.; Goodpaster, J. D. *J. Chem. Theory Comput.* **2018**, *14*, 1928–1942.

(29) Zgid, D.; Chan, G. K. L. *J. Chem. Phys.* **2011**, *134*, 94115.

(30) Zgid, D.; Gull, E.; Chan, G. K. L. *Phys. Rev. B* **2012**, *86*, 165128.

(31) Knizia, G.; Chan, G. K. L. *Phys. Rev. Lett.* **2012**, *109*, 186404.

(32) Knizia, G.; Chan, G. K. L. *J. Chem. Theory Comput.* **2013**, *9*, 1428.

(33) Zhu, T.; Cui, Z.-H.; Chan, G. K.-L. *J. Chem. Theory Comput.* **2020**, *16*, 141.

(34) Cui, Z.-H.; Zhu, T.; Chan, G. K.-L. *J. Chem. Theory Comput.* **2020**, *16*, 119–129.

(35) Masur, O.; Schütz, M.; Maschio, L.; Usvyat, D. *J. Chem. Theory Comput.* **2016**, *12*, 5145–5156.

(36) Schütz, M.; Maschio, L.; Karttunen, A. J.; Usvyat, D. *J. Phys. Chem. Lett.* **2017**, *8*, 1290.

(37) Usvyat, D.; Maschio, L.; Schütz, M. *WIREs: Comput. Mol. Sci.* **2018**, *8*, e1357.

(38) Christlmaier, M.; Kats, D.; Alavi, A.; Usvyat, D. Full configuration interaction quantum Monte Carlo for fragments embedded in periodic mean field (in preparation).

(39) Kats, D.; Manby, F. R. *J. Chem. Phys.* **2013**, *139*, 021102.

(40) Kats, D.; Köhn, A. *J. Chem. Phys.* **2019**, *150*, 151101.

(41) Schütz, M.; Usvyat, D.; Lorenz, M.; Pisani, C.; Maschio, L.; Casassa, S.; Halo, M. *Accurate Condensed-Phase Quantum Chemistry, Series: Computation in Chemistry*; CRC Press, 2010; Vol. 27; Chapter Density fitting for correlated calculations in periodic systems.

(42) Usvyat, D. *J. Chem. Phys.* **2013**, *139*, 194101.

(43) Dovesi, R.; Erba, A.; Orlando, R.; Zicovich-Wilson, C. M.; Civalleri, B.; Maschio, L.; Rérat, M.; Casassa, S.; Baima, J.; Salustro, S.; Kirtman, B. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2018**, *8*, e1360.

(44) Pisani, C.; Schütz, M.; Casassa, S.; Usvyat, D.; Maschio, L.; Lorenz, M.; Erba, A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7615–7628.

(45) Shao, Y. et al. *Mol. Phys.* **2014**, *113*, 184–215.

(46) Zicovich-Wilson, C. M.; Dovesi, R.; Saunders, V. R. *J. Chem. Phys.* **2001**, *115*, 9708–9719.

(47) Usvyat, D.; Maschio, L.; Pisani, C.; Schütz, M. *Zeitschrift fur Phys. Chemie* **2010**, *224*, 441–454.

(48) Knowles, P. J.; Handy, N. C. *Comput. Phys. Commun.* **1989**, *54*, 75–83.

(49) Casanova, D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9779–9790.

(50) Casanova, D. *J. Chem. Phys.* **2012**, *137*.

(51) Zimmerman, P. M.; Bell, F.; Goldey, M.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2012**, *137*.

(52) Casanova, D. *J. Comput. Chem.* **2013**, *34*, 720–730.

(53) Bell, F.; Zimmerman, P. M.; Casanova, D.; Goldey, M.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 358–366.

(54) Casanova, D. *J. Chem. Phys.* **2014**, *140*.

(55) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1988**, *89*, 2185–2192.

(56) Pipek, J.; Bogár, F. *Topics in Current Chemistry*; Springer Berlin Heidelberg, 1999; pp 43–61.

(57) Sousa, C.; Domínguez-Ariza, D.; De Graaf, C.; Illas, F. *J. Chem. Phys.* **2000**, *113*, 9940–9947.

(58) Werner, H.-J. et al. *J. Chem. Phys.* **2020**, *152*, 144107.

(59) Hanauer, M.; Köhn, A. *J. Chem. Phys.* **2011**, *134*, 204111.

(60) Kallay, M. et al. *J. Chem. Phys.* **2020**, *152*, 074107.

(61) Mayhall, N. J.; Head-Gordon, M. *J. Phys. Chem. Lett.* **2015**, *6*, 1982–1988.

(62) Chien, A. D.; Zimmerman, P. M. *J. Chem. Phys.* **2017**, *146*, 014103.

(63) Meitei, O. R.; Houck, S. E.; Mayhall, N. J. *J. Chem. Theory Comput.* **2020**, *16*, 3597–3606.

(64) Pham, H. Q.; Bernales, V.; Gagliardi, L. *J. Chem. Theory Comput.* **2018**, *14*, 1960.

(65) Hermes, M. R.; Gagliardi, L. *J. Chem. Theory Comput.* **2019**, *15*, 972.

(66) He, N.; Evangelista, F. *J. Chem. Phys.* **2020**, *152*, 094107.

Graphical TOC Entry