



Molecular Physics

An International Journal at the Interface Between Chemistry and Physics

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/tmph20>

Smooth potential-energy surfaces in fragmentation-based local correlation methods for periodic systems

A. S. Hansen, E. Aurbakken & T. B. Pedersen

To cite this article: A. S. Hansen, E. Aurbakken & T. B. Pedersen (2021): Smooth potential-energy surfaces in fragmentation-based local correlation methods for periodic systems, *Molecular Physics*, DOI: [10.1080/00268976.2021.1896046](https://doi.org/10.1080/00268976.2021.1896046)

To link to this article: <https://doi.org/10.1080/00268976.2021.1896046>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 04 Mar 2021.



[Submit your article to this journal](#)



Article views: 232



[View related articles](#)



[View Crossmark data](#)

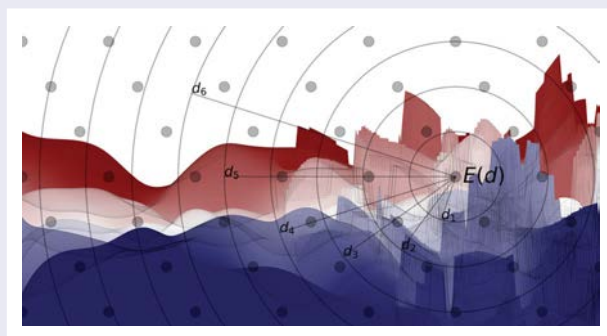
Smooth potential-energy surfaces in fragmentation-based local correlation methods for periodic systems

A. S. Hansen , E. Aurbakken and T. B. Pedersen 

Department of Chemistry, Hylleraas Centre for Quantum Molecular Sciences, University of Oslo, Oslo, Norway

ABSTRACT

Local approximations facilitate the application of post-Hartree–Fock methods in the condensed phase, but simultaneously introduce errors leading to discontinuous potential-energy surfaces. In this work, we explore how these discontinuities arise in periodic systems, their implications, and possible ways of controlling them. In addition, we present a fully periodic Divide-Expand-Consolidate second-order Møller–Plesset approach using an attenuated resolution-of-the-identity approximation for the electron repulsion integrals and a convenient class to handle translation-symmetric tensors in block-Toeplitz format.



ARTICLE HISTORY

Received 16 December 2020
Accepted 19 February 2021

KEYWORDS

Periodic systems; MP2; attenuated fitting; local correlation; fragmentation



1. Introduction

The theoretical description of electronic interactions at the quantum level is fundamentally the same for periodic systems as for molecules. Challenges posed by the computational scaling of correlated wavefunction methods have, however, proven much more limiting in the periodic case, and condensed-phase wavefunction-based results beyond Kohn–Sham density-functional theory (KS-DFT) [1] or Hartree–Fock (HF) [2] theory remain fairly uncommon in the literature relative to the molecular case.

With today's line-up of efficient and publicly available coupled-cluster (CC) implementations, it is possible to perform highly accurate electronic structure calculations for moderately sized molecules at an affordable computational cost [3]. While valuable in their own right, such

results are often used to ensure the quality of other less reliable but more efficient methods, such as KS-DFT or semi-empirical methods. The results provided by these implementations would have been unattainable, if not for several important developments made in the past 2–3 decades, where the computational scaling with respect to the number of electrons has been reduced towards linearity by means of various approximations. Distance-based approximations dating back to Pulay and others [4–10] commonly referred to as *local correlation methods* have been essential in this regard.

Notably, we have seen the revival of Meyer's pair natural orbitals (PNO) [11] in the work from Neese [12–16], followed by Werner [17], where the most significant part of the virtual orbital space is represented in a compact manner for each pair of occupied orbitals.

CONTACT A. S. Hansen  a.s.hansen@kjemi.uio.no  Department of Chemistry, Hylleraas Centre for Quantum Molecular Sciences, University of Oslo, P.O. Box 1033 Blindern, Oslo N-0315, Norway.

 Supplemental data for this article can be accessed here. <https://doi.org/10.1080/00268976.2021.1896046>

Simultaneously, extensive research into the decoupling of the orbital spaces has been made. Various approaches for decoupling the CC equations have been presented in the literature, including the divide-expand-consolidate (DEC) approach by Jørgensen and coworkers [18–22] and the closely related cluster-in-molecule (CIM) approach by Li et al. [23,24], and the generalisation of the Pulay–Sæbø scheme by Werner and coworkers [25], to name a few.

While these methods have greatly reduced the computational cost of molecular CC calculations, their adaptation to the periodic case has not been without complications. The natural way of dealing with the electronic wavefunction in bulk materials is to assume an infinite periodic extent, which subsequently allows mapping the infinite problem onto the finite unit cell by means of a Fourier transform [26,27]. The solution of these equations, the Bloch orbitals, are expressed as linear combinations of delocalised plane waves [28]. As a consequence, periodic systems are not just simply huge molecules – they are truly infinite, yet subject to symmetries which reduce the computational expense of the correlation treatment. One approach to correlation methods in solids thus is to deal with the problem in its canonical form, expanded in either plane waves [29–31] or localised atomic orbitals such as Gaussian basis functions [32–36]. The connection to the more familiar local quantum chemical treatment of correlation is, however, conditioned on the availability of a Wannier representation of the orbital spaces [37], which is restricted to non-conducting systems [38–41]. In this representation, each orbital is no longer uniquely associated with a wave vector and the decoupling conditions are thus different than in the Bloch case. Still, the prospect of achieving similar results as for molecules has motivated a line of research into the application of these methods on periodic systems, as documented by the work of Usvyat et al. [42], Li et al. [43], and ourselves [44,45]. Beyond alleviating the notable computational challenges of obtaining the ground-state solution, local correlation methods may provide novel pathways towards the calculation of band gaps, thus providing promising venues for application and a closer connection to experiment [46,47].

As we gain experience with these methods in the periodic realm, it becomes increasingly clear that an efficient implementation requires many layers of approximations, rather than one single principle. This includes approximate schemes for the electron repulsion integrals (ERIs) [48–50] and screening, as well as the local approximation itself. While the resolution of the identity (RI) approximation has been demonstrated to greatly accelerate ERI generation and transformation in local second-order Møller-Plesset (MP2) theory calculations [51,52],

there are open questions regarding fitting basis sets and methodology which warrants further studies [50,53]. From new layers of approximations, new input parameters follow, each affecting the final result of the calculation in various ways. Careful consideration should thus be taken in order to retain the best possible control of the error in the result, so it can be suppressed or extrapolated away in a systematic manner. In this work, we focus on the error due to truncated orbital spaces, with special focus on how it impacts potential-energy surfaces (PESs) at the MP2 level of theory for periodic systems. This is likely the most significant error present in these calculations [54–57] and should be treated accordingly.

2. Theory

Within the adiabatic Born–Oppenheimer approximation, the electronic ground-state energy of a periodic many-electron system defines the PES as a function of the parameters describing the lattice and the coordinates of the atoms associated with one unit cell, the reference cell, in the lattice. The accuracy of the PES depends on the level of theory used to calculate the energy, and can in principle be treated systematically within convergent hierarchies of wavefunction methods, such as the CC methods [3]. However, the steep computational scaling of post-HF methods combined with the infinite nature of periodic systems impose severe limitations already at the level of MP2 theory, which has resulted in a considerable effort being put into the adaptation of local correlation methods [4–10] to the periodic case. While these methods greatly reduce the computational cost of post-HF calculations, they are known to simultaneously introduce discontinuities in PESs which can potentially be difficult to control [54,58].

In order to understand how these discontinuities emerge in periodic systems, we shall look closer into how the local energies are computed in the first place. The periodic Hartree–Fock equations are conventionally solved in reciprocal space using either a plane-wave basis or a translational symmetry-adapted local atomic orbital (AO) basis such as Gaussian AOs. In both cases [59], the resulting canonical representation of the orbital space is the completely delocalised and complex-valued Bloch orbitals [28].

For insulators, it is possible to obtain a non-canonical, local representation of the orbital space by means of an inverse Fourier transform and localisation in direct space [41]. In this picture, the space is spanned by Wannier orbitals φ_{Lp} , where the uppercase, boldfont index refers to the cell in the periodic structure in which the orbital p belongs.

In addition to being local in the sense of having a finite spatial extent, the Wannier orbitals are translationally orthogonal to each other [41]:

$$\int_{\mathbb{R}^3} \varphi_{Lp} \varphi_{Mq} d\mathbf{r} = \delta_{LM} \delta_{pq}, \quad (1)$$

and can be real-valued by constraining the localisation procedure [41].

Post-HF calculations from a non-canonical reference determinant must account for non-diagonal terms in the Fock matrix, meaning that, for instance, computing the MP2 correlation energy is no longer a one-step procedure. Pulay and Sæbø [9], therefore, derived an MP2 formulation from the Hylleraas functional that is independent of unitary rotations within the occupied and virtual spaces. In the periodic case, for Wannier functions, these equations are

$$0 = (\mathbf{I}i\mathbf{A}a|\mathbf{J}j\mathbf{B}b) + \sum_{Cc} t_{\mathbf{I}ij}^{CcBb} f_{\mathbf{A}aCc} + \sum_{Cc} t_{\mathbf{I}ij}^{AaCc} f_{\mathbf{B}bCc} - \sum_{Kk} t_{\mathbf{K}kj}^{AaBb} f_{\mathbf{K}kIi} - \sum_{Kk} t_{\mathbf{I}Kk}^{AaBb} f_{\mathbf{K}kIj}. \quad (2)$$

where f is the Fock matrix, t are the MP2 amplitudes and the periodic ERIs for real Wannier functions are

$$(\mathbf{0}i\mathbf{A}a|\mathbf{J}j\mathbf{B}b) := \iint_{\mathbb{R}^3} \frac{\varphi_{\mathbf{0}i}(\mathbf{r}) \varphi_{\mathbf{A}a}(\mathbf{r}) \varphi_{\mathbf{J}j}(\mathbf{r}') \varphi_{\mathbf{B}b}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' d\mathbf{r}. \quad (3)$$

The cell summations in Equation (2) run over the full, infinite lattice. Due to the translational symmetry in the amplitudes and ERIs, given by

$$t_{\mathbf{I}ij}^{AaBb} = t_{\mathbf{0}i(\mathbf{J}-\mathbf{I})j}^{(\mathbf{A}-\mathbf{I})a(\mathbf{B}-\mathbf{I})b}, \quad (4)$$

we may choose to keep the first occupied index fixed in the reference cell $\mathbf{0}$ without any loss of information. Upon solving Equation (2) for the amplitudes, the correlation energy per unit cell can be computed as

$$\Delta E_{\text{MP2}} = \sum_{\mathbf{J}} \sum_{\mathbf{A}\mathbf{B}} \sum_{ij} \sum_{ab} t_{\mathbf{0}ij}^{AaBb} (2 (\mathbf{0}i\mathbf{A}a|\mathbf{J}j\mathbf{B}b) - (\mathbf{0}i\mathbf{B}b|\mathbf{J}j\mathbf{A}a)). \quad (5)$$

It is well known that for localised orbitals, the ERIs tend exponentially to zero with the *inter-orbital* distances $R_{\mathbf{0}i\mathbf{A}a}$ and $R_{\mathbf{J}j\mathbf{B}a}$, and proportional to $R_{\mathbf{0}ij}^{-3}$ with the *inter-pair* distance $R_{\mathbf{0}ij}$ [60,61], where

$$R_{\mathbf{P}p\mathbf{Q}q} = \sqrt{(\langle \varphi_{\mathbf{P}p} | \hat{\mathbf{r}} | \varphi_{\mathbf{P}p} \rangle - \langle \varphi_{\mathbf{Q}q} | \hat{\mathbf{r}} | \varphi_{\mathbf{Q}q} \rangle)^2}. \quad (6)$$

These decay properties reveal a certain decoupling in the amplitude equations and form the basis for various local

correlation methods. Linear scaling can be achieved by partitioning the summations in Equations (2) and (5) into weakly coupled subspaces Ω_k , where only the subsets of occupied and virtual orbitals that significantly affect the energy of these subspaces are included. After solving the equations on these fragmented subspaces, the total correlation energy can finally be approximately expressed as a sum over the contribution from each:

$$\Delta E_{\text{MP2}} \approx \sum_k \Delta E_{\text{MP2}}(\Omega_k). \quad (7)$$

The equations are typically solved on slightly larger domains by inclusion of so-called buffer spaces [18] to ensure that the energy inside each subspace is properly converged.

The correlation energy is distributed across the excitation space, and will smoothly flow between configurations following smooth changes in geometry and representations of the orbital spaces. Changes in the weakly coupled local subspaces following the inclusion or exclusion of orbitals depending on distance measures will, however, inevitably introduce discontinuous changes in the energy. Consequently, local correlation methods are prone to yield non-smooth – or *fractured* – PESs and careful consideration should thus be put into the algorithms by which these local subspaces are constructed in order to minimise such effects.

We shall distinguish the fragmentation schemes, i.e. the principles used to construct the subspaces, from the various other approximations involved. In terms of fragmentation, the challenges associated with fractured PESs have been addressed in various ways, with the most common being that of choosing the subspaces sufficiently large to suppress discontinuities [57]. Other approaches include freezing the domains close to the equilibrium geometry [55] or bumping the amplitudes [56]. The orbital specific virtual (OSV) representation of the virtual space has been of particular interest, as it has been shown to yield smoother surfaces both for molecules and periodic systems [57,62–64].

The divide-expand-consolidate (DEC) family of methods takes a different approach and aims at controlling the magnitude of the discontinuities by converging the energy with respect to changes in the orbital subspaces. The various fragmentation schemes are themselves in principle agnostic with respect to the representations of the orbital spaces, which for the virtual space include local virtual orbitals (LVOs), projected atomic orbitals (PAOs), OSVs and pair natural orbitals (PNOs) [11–17]. While it is well known that the choice of virtual representation makes a significant impact on the energy and thus the PES, it currently remains unclear

how to determine the suitability of such a representation without performing the actual energy calculation [45,65].

Positions and locality can be ascribed to a set of orbitals in a number of ways [45], and the exact definition of the local subspaces on which the amplitude equations are solved may therefore differ between various implementations. Following an “ERI-centric” approach, we shall infer sparsity in the ERIs directly from distance considerations, meaning that we assume $(\mathbf{O}i\mathbf{A}a|Jj\mathbf{B}b) = 0$ for distances $R_{\mathbf{O}i\mathbf{A}a}$, $R_{Jj\mathbf{B}b}$ and $R_{\mathbf{O}iJj}$ above some predetermined thresholds. This approach is convenient from a periodic perspective, since it incorporates the translational symmetry of the product orbitals that appear in the bra and ket of the ERIs,

$$\hat{T}_{\mathbf{M}}\varphi_{\mathbf{O}i}(\mathbf{r})\varphi_{\mathbf{A}a}(\mathbf{r}) = \varphi_{\mathbf{M}i}(\mathbf{r})\varphi_{(\mathbf{A}+\mathbf{M})a}(\mathbf{r}), \quad (8)$$

where $\hat{T}_{\mathbf{M}}$ is the lattice translation operator that offsets the product orbitals in cellwise increments, such that any excitation included in the reference cell is also included in the translated copies.

By defining the elements

$$E_{Iij}^{\mathbf{A}a\mathbf{B}b} := t_{Iij}^{\mathbf{A}a\mathbf{B}b} (2(\mathbf{I}i\mathbf{A}a|Jj\mathbf{B}b) - (\mathbf{I}i\mathbf{B}b|Jj\mathbf{A}a)), \quad (9)$$

we may express the full MP2 energy per cell more compactly as

$$\Delta E_{\text{MP2}} = \sum_{i \in \mathcal{O}_0} \sum_{Jj \in \mathcal{O}} \sum_{\mathbf{A}a \in \mathcal{V}} \sum_{\mathbf{B}b \in \mathcal{V}} E_{\mathbf{O}iJj}^{\mathbf{A}a\mathbf{B}b}, \quad (10)$$

where the summations are now partitioned in such a way that the individual orbitals are emphasised rather than the lattice structure. The summation is still infinite, but we explicitly denote the full orbital spaces as \mathcal{V} and \mathcal{O} (\mathcal{O}_0 for the reference cell) for the virtual and occupied spaces, respectively.

In this setting, a *fragmentation* may be regarded as a partitioning within the various terms, while the local approximation may be regarded as a distance-based truncation of the orbital spaces. Taken together, this forms a *fragmentation scheme*, and with the addition of approximative techniques such as various representations of the orbital spaces, it can be seen as a *local correlation method*, as found in a number of implementations.

The DEC family of methods [18–20] was originally devised for molecules and has been extended to the periodic case recently [44]. In the DEC approach, the amplitude equations are solved within subspaces called amplitude orbital spaces (AOSs) and the resulting energies are computed on a smaller subspace of each AOS called the energy orbital space (EOS). In order to gain control over the error in these calculations, these spaces are constructed such that the error in the energy is below a predefined threshold.

In terms of fragmentation, the DEC approach first divides the occupied space into fragments, which are non-overlapping sets \underline{P} of occupied orbitals in each others vicinity. The AOS for a fragment is

$$\Omega_{k(\mathbf{O}P)} = \mathcal{O}_{\underline{P}} \cup \mathcal{V}_{\mathbf{O}P}, \quad (11)$$

where the notation $\mathcal{O}_{\underline{P}}$ and $\mathcal{V}_{\mathbf{O}P}$ signifies sets of occupied and virtual orbitals in the vicinity of \underline{P} (explicit definition will follow later). The energy associated with a fragment is in terms of Equation (7):

$$\Delta E_{\text{frag,MP2}}(\Omega_{k(\mathbf{O}P)}) = \sum_{ij \in \underline{P}} \sum_{\mathbf{A}a\mathbf{B}b \in \mathcal{V}_{\mathbf{O}P}} E_{\mathbf{O}i\mathbf{O}j}^{\mathbf{A}a\mathbf{B}b}, \quad (12)$$

where the summation domain constitutes the EOS of the fragment. The AOS of a fragment can be determined by expanding $\mathcal{O}_{\underline{P}}$ and $\mathcal{V}_{\mathbf{O}P}$ successively until the change in energy is below a certain threshold referred to as the fragment optimisation threshold (FOT). The converged fragment spaces obtained in the expansive step may then be used to set up the so-called pair-fragment spaces with AOSs

$$\Omega_{k(\mathbf{O}P, \mathbf{L}Q)} = \mathcal{O}_{\underline{P}} \cup \mathcal{V}_{\mathbf{O}P} \cup \mathcal{O}_{\underline{L}Q} \cup \mathcal{V}_{\mathbf{L}Q}, \quad (13)$$

on which the pair-fragment amplitude equations are then solved. Finally, the fragment energies can be consolidated with the pair-fragment energies of successively increasing inter-pair distance until the energy is converged, where the pair-fragment energy is

$$\Delta E_{\text{pair,MP2}}(\Omega_{k(\mathbf{O}PLQ)}) = \sum_{i \in \underline{P}} \sum_{j \in \underline{Q}} \sum_{\mathbf{A}a\mathbf{B}b \in \mathcal{V}_{\mathbf{O}P} \cup \mathcal{V}_{\mathbf{L}Q}} E_{\mathbf{O}i\mathbf{L}j}^{\mathbf{A}a\mathbf{B}b}, \quad (14)$$

where again the summation domain is the EOS of the pair-fragment. In contrast to molecular DEC we have in the above made no distinction between the virtual spaces used in the AOS and EOS. The AOS does, however, contain buffer orbitals in the form of occupied orbitals used merely to converge the energy inside the EOS.

The partitioning used to arrive at the sets \underline{P} can of course be varied, all the way from no partitioning ($\underline{P} = \mathcal{O}_0$) to full partitioning ($\underline{P}_i = \varphi_i$). In the latter case, this closely resembles what is known as the Pulay–Sæbø approach [7,66], where every unique combination of occupied orbitals are referred to as *pairs*. The fragmentation, which has later been generalised by Werner and coworkers [25,67], has been extended to periodic systems in the CRYSCOR program [42,51,57,68–71]. In this approach, the DEC fragment is referred to as a *strong pair*, while the pair fragments are classified into *close*, *weak* or *distant* pairs depending on the inter-orbital distance

between the constituent occupied orbitals. In the context of Equation (7), these subspaces can be expressed as

$$\Omega_{k(0\mathbf{L}j)} = \varphi_{0i} \cup \mathcal{V}_{0i} \cup \varphi_{\mathbf{L}j} \cup \mathcal{V}_{\mathbf{L}j}, \forall \mathbf{L} \in \mathbb{Z}^3, \quad (15)$$

and the energy of such a subspace is

$$\Delta E_{\text{MP2}}(\Omega_{k(0\mathbf{L}j)}) = \sum_{AaBb \in \mathcal{V}_{0i} \cup \mathcal{V}_{\mathbf{L}j}} E_{0\mathbf{L}j}^{AaBb}. \quad (16)$$

The pair-classification scheme is well suited for a multi-level approach, where the pairs can be treated within various approximations, ranging from exact MP2 for strong and close pairs, to multipole expansion for weak pairs and Lennard–Jones extrapolation for distant pairs. Its implementation in `CRYSCOR` is currently capable of running MP2-calculations on systems with more than 100 atoms per cell. [49]

With the inclusion of the occupied buffer space in the vicinity of a given fragment (or strong pair), it is tempting to include the corresponding amplitudes in the energy calculation since they are likely already close to convergence. This can be done by extending each fragment (or strong pair) space $\underline{P}_i = \varphi_i$ with all pair-fragment spaces where one of the occupied orbitals pertain to the given fragment, forming the subspace

$$\Omega_{k(0i)} = \varphi_{0i} \cup \mathcal{O}_{0i} \cup \mathcal{V}_{0i}, \quad (17)$$

and computing the energy

$$\Delta E_{\text{MP2}}(\Omega_{k(0i)}) = \sum_{\mathbf{J}j \in \mathcal{O}_{0i}} \sum_{AaBb \in \mathcal{V}_{0i}} E_{0\mathbf{L}j}^{AaBb}. \quad (18)$$

This fragmentation corresponds to the periodic cluster-in-molecule (CIM) method [43], where the constituent subspaces are referred to as *clusters*. In CIM, the spaces are determined by including all orbitals up to a given distance, typically fixed at 5.5 Å [43].

To summarise, the DEC and Pulay–Sæbø fragmentation schemes become identical in the case where each DEC-fragment is comprised of one single occupied orbital, while the CIM fragmentation differs from the two others in the sense that the pair-contributions are always included in each cluster. An illustrative comparison of the three fragmentation schemes is presented in Figure 1. While the smoothness of the PES in the DEC approach can in principle be controlled by the FOT, discontinuities in the Pulay–Sæbø or CIM scheme are conditioned on the default or user-defined domains being sufficiently large. As with the incremental scheme DEC offers a systematic path towards the exact correlation energy for solids, yet in smaller increments and applicable only to insulators [72,73]. All three fragmentation

schemes may serve as suitable starting points for embedding approaches, where higher-level correlation methods designed for molecules can be used to improve upon the results [74–76].

It thus turns out that by not fully contracting the energy, but rather solving for the amplitudes within the local subspace and thereafter storing the elements $E_{0\mathbf{L}j}^{AaBb}$ in addition to the positions of the orbitals, the energy can be retrospectively recast as a Pulay–Sæbø-like, CIM-like or DEC-like result. Furthermore, by computing a range of such energies for various local truncations of the occupied and virtual subspaces, it becomes possible to gain insight into how and where discontinuities emerge in the PES depending on the cutoff parameters. We have implemented this functionality in the current XDEC-code, but stress that this procedure does not equate to `CRYSCOR` or CIM, since it lacks the many other approximative techniques and orbital space representations that are used in these implementations.

3. Implementation

The implementation of XDEC presented in this paper shall be referred to as XDEC-RI (extended DEC with a resolution of the identity approximation for the ERIs) in order to distinguish it from our previous implementation, XDEC, since there are fundamental differences between the two.

As with our XDEC implementation [44], the Wannier functions in XDEC-RI are obtained from `CRYSTAL` [41,71] and the AO integrals from `LIBINT` [77]. A notable change is the fact that the ERIs are now fitted to a periodic basis using an attenuated Coulomb metric [50]. A global fitting at the onset of the calculation saves computational time, since the ERIs can be easily computed on demand and reused in the expansive steps of the fragment optimisation. Furthermore, all integrals and orbital spaces now fully incorporate translational symmetry.

The size of the local subspace is controlled as follows. Each WF is assigned a position according to its centre. The occupied orbitals may thereafter be subdivided into fragments depending on their positions or they may be kept in separate fragments depending on the user’s preference. A fragment centre is typically chosen to be the centre of the first orbital in each fragment. The local subspace pertaining to the fragment is then chosen to consist of all occupied and virtual orbitals inside given radii d_{occ} and d_{virt} .

In our previous implementation of the XDEC algorithm [44], we treated fragments and pair-fragments as finite clusters. Although this, in principle, would allow for the exploitation of highly optimised molecular algorithms, we here report an implementation that maintains

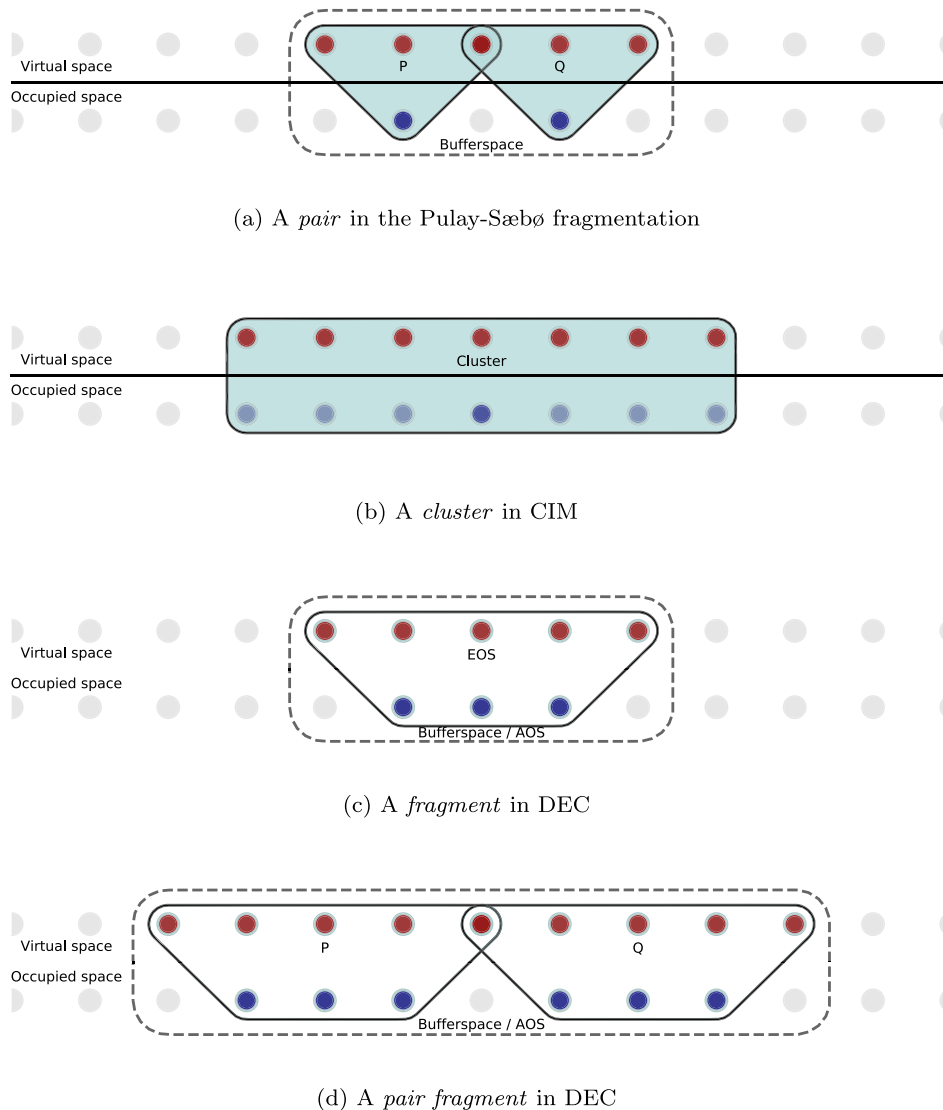


Figure 1. Illustration of the various periodic fragmentation schemes under consideration. Virtual and occupied orbitals inside the local subspaces are coloured red and blue, respectively. CIM differs from the other two in the sense that the energy contributions from amplitudes where i and j are not the same is included in the central cluster of i , here indicated by varying opacity in the occupied orbitals. (a) A *pair* in the Pulay-Sæbø fragmentation. (b) A *cluster* in CIM. (c) A *fragment* in DEC. (d) A *pair fragment* in DEC.

full translational symmetry throughout the calculation. This allows us to circumvent the main bottleneck of the XDEC algorithm, namely the repeated evaluation of ERIs, by reusing integrals in the fragment and pair-fragment calculations. The main disadvantage of this approach is that the resulting algorithm is no longer embarrassingly parallel due to the increased amount of communication between computer nodes.

To support full translational symmetry we have implemented a Python class named *TMAT* (Toeplitz Matrix), which is specifically designed to facilitate linear-algebra operations on bi-infinite block-Toeplitz and block-Circulant matrices [78] (definitions are given below).

Using local Gaussian basis sets from quantum chemistry, translational symmetry manifests itself in

equivalence among blocks of matrix elements of translationally invariant operators. A typical example is the infinite overlap matrix, whose elements obey

$$S_{\mu\nu}^{\mathbf{M},\mathbf{N}} = \langle \mathbf{M}\mu | \mathbf{N}\nu \rangle = S_{\mu\nu}^{\mathbf{M}+\mathbf{L},\mathbf{N}+\mathbf{L}}. \quad (19)$$

Matrices with this block structure are called bi-infinite block Toeplitz (IBT) matrices. Using $i, j \in \mathbb{Z}$ to denote row and column block indices, an IBT matrix is defined by

$$\mathbf{A}^{ij} = \mathbf{A}^{j-i} := \mathbf{A}^m. \quad (20)$$

The matrix product of two IBT matrices \mathbf{A} and \mathbf{B} is itself an IBT matrix, since

$$(\mathbf{AB})^{(i+n),(j+n)} = \sum_{k=-\infty}^{\infty} A^{(k-n)-i} B^{j-(k-n)} = (\mathbf{AB})^{ij}, \quad (21)$$

and can thus be compactly expressed as

$$(\mathbf{AB})_{IBT}^m = \sum_{k=-\infty}^{\infty} A^k \cdot B^{m-k}. \quad (22)$$

If we assign to \mathbf{A} a bandwidth N such that

$$A^m = 0 \text{ for } |m| > N \quad (23)$$

we can define a discrete Fourier transformation

$$\mathcal{F}\{\mathbf{A}\}^m = \sum_{n=-N}^N e^{-i(\frac{2\pi}{M})mn} A^n := \tilde{A}^m, \quad (24)$$

and its inverse

$$\mathcal{F}^{-1}\{\tilde{\mathbf{A}}\}^n = \frac{1}{M} \sum_{m=-N}^N e^{-i(\frac{2\pi}{M})mn} \tilde{A}^m = A^n, \quad (25)$$

where $M = 2N + 1$. The discrete Fourier transformation followed by its inverse introduces a type of periodicity in the matrices that satisfy the Born–von Karman boundary condition [28], as apparent when considering blocks outside the bandwidth:

$$A^{M+n} = \sum_{n'=-N}^N A^{n'} \left(\frac{1}{M} \sum_{m=-N}^N e^{i2\pi \frac{m}{M}(M+n-n')} \right) = A^n. \quad (26)$$

In order to account for this behaviour, it makes sense to define a non-standard *circulant* operator

$$i \circlearrowleft N := (i + N) \bmod (2N + 1) - N, \quad (27)$$

where *mod* implies the remainder of integer division, producing the following infinite series (in positive increments of 1):

$$\dots, N, -N, \dots, 0, \dots, N, -N, \dots, \quad (28)$$

such that the transformation to and from reciprocal space can be expressed

$$A^n = \sum_{n'=-N}^N A^{n'} \underbrace{\left(\frac{1}{M} \sum_{m=-N}^N e^{i2\pi \frac{m}{M}(n-n')} \right)}_{\delta_{n \circlearrowleft N, n'}}, \quad (29)$$

where we explicitly have pointed out the circulant nature of the series expansion of the Kronecker delta by including the circulant operator. A matrix subject to these conditions can be referred to as an *infinite block-circulant*

(IBC) matrix, defined by

$$\mathbf{A}^{ij} = A^{(j-i) \circlearrowleft N} := A^{n \circlearrowleft N}. \quad (30)$$

We now consider for given n the inverse Fourier transform of an element-wise multiplication of two IBT matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ in reciprocal space:

$$\mathcal{F}^{-1}(\mathcal{F}(\mathbf{A}) * \mathcal{F}(\mathbf{B}))^n = \frac{1}{M} \sum_{m=-N}^N e^{i\frac{2\pi}{M}mn} \tilde{A}^m \tilde{B}^m. \quad (31)$$

Expanding $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ in terms of their blocks in direct space,

$$\begin{aligned} \mathcal{F}^{-1}(\mathcal{F}(\mathbf{A}) * \mathcal{F}(\mathbf{B}))^n &= \sum_{n_1=-N}^N \sum_{n_2=-N}^N \\ &\times \left(\frac{1}{M} \sum_{m=-N}^N e^{i\frac{2\pi}{M}m(n-n_1-n_2)} \right) A^{n_1} B^{n_2}, \quad (32) \end{aligned}$$

which by the circulant Kronecker delta can be reduced to the IBC product

$$\begin{aligned} \mathcal{F}^{-1}(\mathcal{F}(\mathbf{A}) * \mathcal{F}(\mathbf{B}))^n &= \sum_{n_1=-N}^N A^{n_1} B^{n-n_1 \circlearrowleft N} \\ &= (\mathbf{AB})_{IBC}^n. \quad (33) \end{aligned}$$

A comparison of Equations (33) and (22) reveals the relation

$$\begin{aligned} (\mathbf{AB})_{IBT}^n &= (\mathbf{AB})_{IBC}^n \\ &- \sum_{k=1}^{k \leq |n|} A^{-\text{sgn}(n)(N-k+1)} B^{n-\text{sgn}(n)(N+k)}, \quad (34) \end{aligned}$$

where the superscript IBT indicates the result from Equation ((22)), thus showing that these products can be made equivalent in any given bandwidth by padding the outer layers in the matrices with zeros so that the last term in the above is zero.

Equation (33) is basically the familiar convolution theorem. It provides key insight into the reciprocal space treatment of infinite systems. In conjunction with Equation ((34)), it shows that any factorisation in reciprocal space carries over to direct space in a well-defined manner, paving the way for a straightforward implementation via reciprocal space of widely used linear-algebra techniques such as singular value decomposition, Cholesky factorisation, or diagonalisation.

Furthermore, with a periodic definition of the matrix product there is no need for explicit inclusion of the

lattice-summations in the notation, as they simply follow from the definition. Most of the familiar tensor notation in quantum chemistry thus smoothly carries over to the periodic domain, making the process of extending methods from finite to extended systems conceptually simpler. By implementing the IBT and IBC matrix structures as a class in Python, the process of writing a periodic code much more closely resembles what is typically done for the molecular case, and the cumbersome process of dealing with summations over lattice vectors is conveniently dealt with automatically in the background. Efficiency is ensured by using the Fast Fourier Transform (FFT) algorithm [79] and level 3 BLAS routines [80], exploiting multi-threaded and architecture-optimised libraries. This computational efficiency, in conjunction with the conditions provided by Equation ((33)), is extensively used in our code.

The calculation of ERIs constitutes the main bottleneck in the XDEC implementation, which requires re-calculation of integrals for each fragment and pair-fragment optimisation, including re-calculation in each optimisation cycle. While embarrassingly parallel [22], this approach is too costly to be applicable on small and medium-sized commodity computer clusters. As mentioned above, we here exploit full translational symmetry to enable the reuse of ERIs across fragments and pair-fragments. Still, the calculation of the ERIs for the MP2 approximation formally scales as N^5 with N being the number of atomic orbitals in the supercell. To accelerate the ERI generation and to reduce the memory footprint of the algorithm, we use a periodic adaptation of the resolution of the identity (RI) approximation [81–91].

In a suitable auxiliary basis $\{\varphi_{\mathbf{NK}}\}$, the product of two Wannier orbitals $\varphi_{\mathbf{0iAa}}$ can be approximated as a linear combination [42]:

$$|\mathbf{0iAa}\rangle \approx \sum_{\mathbf{NK}} |\mathbf{NK}\rangle d_{\mathbf{K},\mathbf{0iAa}}^{\mathbf{N}} := |\widetilde{\mathbf{0iAa}}\rangle. \quad (35)$$

The optimal fitting coefficients d are obtained by minimising the residual norm in any positive, semidefinite metric [86],

$$(\mathbf{0iAa} - \widetilde{\mathbf{0iAa}}|\mathbf{0iAa} - \widetilde{\mathbf{0iAa}}) := (\Delta\mathbf{0iAa}|\Delta\mathbf{0iAa}), \quad (36)$$

upon which the residual is orthogonal to the fitting basis, so that

$$(\mathbf{NK}|\Delta\mathbf{0iAa}) = 0. \quad (37)$$

The latter expression yields a set of fitting equations for d :

$$(\mathbf{NK}|\mathbf{0iAa}) = \sum_{\mathbf{MJ}} (\mathbf{NK}|\mathbf{MJ}) d_{\mathbf{j},\mathbf{0iAa}}^{\mathbf{M}}. \quad (38)$$

If we define the infinite block-Toeplitz matrices \mathbf{O} and \mathbf{V} with elements $O_{\mathbf{K},\mathbf{0iAa}}^{-\mathbf{N}} = (\mathbf{NK}|\mathbf{0iAa})$ and $V_{\mathbf{KJ}}^{\mathbf{M}-\mathbf{N}} :=$

$(\mathbf{NK}|\mathbf{MJ})$, the above relation may be cast into matrix form with a straightforward solution for the coefficients

$$\mathbf{d} = \mathbf{V}^{-1}\mathbf{O}, \quad (39)$$

provided the matrix \mathbf{V} is nonsingular – i.e. – provided the auxiliary basis functions constitute a linearly independent set in the chosen metric.

A natural choice for metric in Equation (37) is the Coulomb operator, as this is the metric of the ERIs themselves. The Coulomb operator, however, decays slowly with distance (R^{-1}), making it computationally expensive for infinite periodic systems. Noting that the Coulomb operator is obtained from an *attenuated* Coulomb operator expressed in terms of the complementary error function [92],

$$\frac{1}{r} = \lim_{\omega \rightarrow 0} \frac{\text{erfc}(\omega r)}{r}, \quad (40)$$

we may enforce more rapid decay by selecting a small positive value for the attenuation parameter ω . The attenuated Coulomb operator maintains characteristics similar to the true Coulomb operator for small ω , while for increasingly large values it approaches the overlap metric. As noted by several authors in the past [48,50,51], this makes it an interesting candidate for fitting metric in periodic systems. We shall indicate the usage of this operator with a tilde in the affected matrices. The fitting equations in the attenuated Coulomb metric thus become

$$\widetilde{\mathbf{V}}\mathbf{d} = \widetilde{\mathbf{O}}. \quad (41)$$

Although the matrices involved are in principle infinite, they exhibit to a certain degree a regular blockwise decay in the elements with respect to the distance to the reference cell which can be tuned by ω . The matrices can therefore be computed up to a given truncation threshold by incrementally including chunks of blocks in spherical shells extending outwards from the reference cell (see Algorithms 1 and 2). We shall use this approach in the construction of the matrices required. The expression for the three-index tensor elements can be reorganised as

$$\begin{aligned} (\mathbf{LJ}|\widetilde{\mathbf{0iAa}}) &= \sum_{\mathbf{M}\mu\mathbf{N}\nu} (\mathbf{LJ}|\widetilde{\mathbf{M}\mu}(\mathbf{M}+\mathbf{N})\nu) c_{\mu i}^{-\mathbf{M}} c_{\nu a}^{-(\mathbf{M}+\mathbf{N})+\mathbf{A}} \\ &:= \sum_{\mathbf{N}\nu} \widetilde{O}_{\mathbf{j}\mathbf{N}\nu}^{-\mathbf{L}} c_{\nu a}^{-\mathbf{N}+\mathbf{A}}, \end{aligned} \quad (42)$$

where c denotes the expansion coefficients of the Wannier orbitals in the AO basis, and

$$\widetilde{O}_{\mathbf{j}\mathbf{N}\nu}^{-\mathbf{L}} := \sum_{\mathbf{M}\mu} (\mathbf{LJ}|\widetilde{\mathbf{M}\mu}\mathbf{N}\nu) c_{\mu i}^{-\mathbf{M}}, \quad (43)$$

is an intermediate contraction tensor for the occupied orbitals.

There are two attractive features of this fitting scheme that makes it especially well suited for XDEC. First, by organising the elements in the intermediate tensor $\tilde{\mathbf{O}}$ in the appropriate way, the contraction of any subset of the virtual orbitals may be done efficiently by means of a matrix product. This is useful for incrementally including larger virtual spaces. Second, the circulant product makes it possible to simultaneously compute for all \mathbf{J}

$$\begin{aligned} (\mathbf{0}i\mathbf{A}a|\mathbf{J}\mathbf{j}(\mathbf{J} + \mathbf{B})b) &\approx (\mathbf{0}\tilde{i}\mathbf{A}a|\mathbf{J}\mathbf{j}(\mathbf{J} + \mathbf{B})b) \\ &= (\tilde{\mathbf{d}}_{\mathbf{0},\mathbf{A}}^T \mathbf{V} \tilde{\mathbf{d}}_{\mathbf{0},\mathbf{B}})^{\mathbf{J}}_{iajb}, \end{aligned} \quad (44)$$

by means of the IBC matrix product in Equation (33). The final contractions are with the full Coulomb matrix, but since the attenuated fitting coefficients are only non-zero on a finite extent of the infinite lattice, we only have to compute the blocks in \mathbf{V} that are required to make the circulant product consistent with the Toeplitz product inside the supercell. For simultaneous excitations inside and outside the supercell (\mathbf{J} outside) we can compute the required extra layers in \mathbf{V} on demand.

While the periodic auxiliary basis is infinite, two additional screening parameters ξ_0 and ξ_1 ensure a finite bandwidth of the RI tensors, as outlined in Algorithms 1 and 2. The screening parameters are in effect a truncation threshold on the level of blocks for the three-index AO integrals and the contracted three-index tensor $\tilde{\mathbf{O}}$. Assuming monotonic decay, the screening has been implemented such that the maximum element of all blocks outside a spherical volume centred on the reference cell is below the threshold. This allows the matrices to be incrementally constructed outwards from the reference cell until the requested precision is reached. We finally remark that this fitting approach is not robust [93]. In order to make it robust, one would have to compute the corresponding \mathbf{O} tensors in Coulomb-metric, defeating the purpose of using attenuation in the first place. Our procedure is specifically designed with flexible orbitals spaces in mind and will likely benefit from incorporating aspects from other similar approaches in the future [53,94].

4. Computational details

The systems studied in this work are chosen to be sufficiently complex for a meaningful analysis, while small enough to make calculations relatively inexpensive and keeping the analysis simple. Similarly, focusing on sources of PES discontinuities rather than on highly accurate results, we use basis sets relatively far from the basis set limit.

Algorithm 1 AO screening procedure implemented in the XDEC-RI-LMP2 code

- 1: Compile a list Ω containing chunks of coordinate vectors $\{\mathbf{M}\}_{\mathbf{R}}$ grouped together into concentric spherical shells in order of increasing radial distance $\|\mathbf{R}_{\mathbf{M}}\|$ to the reference cell.
 - 2: **for** $\{\mathbf{M}\}_{\mathbf{R}}$ in Ω **do**
 - 3: Compute all $(\mathbf{0}\mathbf{J}|\mathbf{0}\mu\mathbf{M}\nu)$'s for the shell
 - 4: **if** all $|(\mathbf{0}\mathbf{J}|\mathbf{0}\mu\mathbf{M}\nu)|_{\max} \leq \xi_0$ **then**
 - 5: Break
 - 6: **end if**
 - 7: Append all \mathbf{M} s to the screened domain Ξ
 - 8: **end for**
 - 9: Set an initial reasonably large cutoff R_{cut}
 - 10: **for** \mathbf{M} in Ξ **do**
 - 11: Compute all cells $(\mathbf{N}\mathbf{J}|\mathbf{0}\mu\mathbf{M}\nu)$ within $R_{\mathbf{N}} \leq R_{\text{cut}}$
 - 12: **if** any $(\mathbf{K}_{\text{boundary}}\mathbf{J}|\mathbf{0}\mu\mathbf{M}\nu) \geq \xi_0$, where $0.95R_{\text{cut}} < R_{\mathbf{K}_{\text{boundary}}} \leq R_{\text{cut}}$ **then**
 - 13: Halt execution, warn/advise user to increase domains.
 - 14: **else**
 - 15: Append all blocks where $(\mathbf{N}\mathbf{J}|\mathbf{0}\mu\mathbf{M}\nu) \geq \xi_0$ to screening domain $\Xi_{\mathbf{M}}$
 - 16: Let $R_{\text{cut}} = 1.1R_{\mathbf{N}_{\text{outer}}}$ where $R_{\mathbf{N}_{\text{outer}}}$ corresponds to the outermost cell in the appended blocks.
 - 17: **end if**
 - 18: **end for**
-

Algorithm 2 MO fitting screening procedure implemented in the XDEC-RI-LMP2 code

- 1: Set $R_{\text{tolerance}} = 10^{-12}$
 - 2: Construct a list Ω of cell-indices \mathbf{L} in order of increasing distance $\|\mathbf{R}_{\mathbf{L}}\|$ to the reference cell.
 - 3: **for** \mathbf{L} in Ω **do**
 - 4: Compute all $\tilde{\mathbf{O}}_{\mathbf{J}\nu\rho}^{\mathbf{L},\mathbf{N}}$
 - 5: **if** $R_{\mathbf{L}} - R_{\mathbf{L}_{\text{prev}}} \geq R_{\text{tolerance}}$ and $|\tilde{\mathbf{O}}_{\mathbf{J}\nu\rho}^{\mathbf{L},\mathbf{N}}| \leq \xi_1$ **then**
 - 6: Break
 - 7: **end if**
 - 8: Store columns $(\mathbf{N}\nu)$ and column indices of $\tilde{\mathbf{O}}_{\mathbf{L}\mathbf{J}\rho,\mathbf{N}\nu}$ with max absolute value above ξ_1 for the subsequent contraction of virtual coefficients.
 - 9: **end for**
-

Validation of the implementation, including selection of the Coulomb attenuation parameter for the RI approximation, is performed using the same three-dimensional (3D) and one-dimensional (1D) systems of neon atoms, and the 1D system of ethylene molecules as in Ref. [44]. In addition, we use a 1D system with a unit cell containing two water molecules arranged as shown in Figure 2.

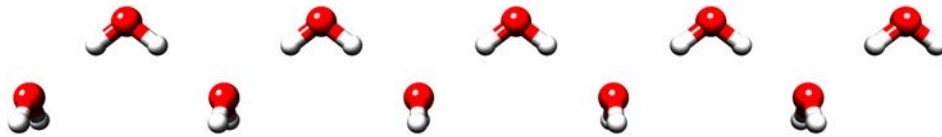


Figure 2. The $2\text{H}_2\text{O}$ system, where two water molecules per cell are repeated in a 1D lattice.

We use an experimentally determined internal geometry of water [95]. For the fixed geometry calculations, we use a lattice parameter of approximately 4.8035 Å (see supplementary information for details). In the PES calculations, we vary the lattice parameter while the O–O–O angles remain fixed. The water system contains more significant inter-molecular correlation effects than ethylene, and is thus more interesting with regard to the exploration of PESs.

Convergence of the MP2 correlation energy with respect to truncation of the local subspaces is investigated for the water and ethylene systems along with water molecules adsorbed on a LiH (001) surface in the slab approximation. The geometry of the water adsorption system is taken from Tsatsoulis et al. [96], with the unit cell containing eight LiH and one water molecule, as shown in Figure 3. This system will be referred to as LiH- H_2O .

Finally, the water and LiH- H_2O systems are used to investigate the continuity of PESs. For the LiH- H_2O system, we start from the equilibrium geometry reported by Tsatsoulis et al. [96] and vary the distance between Li and O to generate a one-dimensional cut through the PES. This system has a lattice parameter of approximately 5.78 Å, thus increasing the interaction between neighbouring water molecules at the surface as compared to systems with lower density. Detailed geometric data for all systems can be found in the supplementary information.

We use the CRYSTAL program [71] to compute the HF reference determinant and to perform the Wannierization and Foster-Boys localisation [41] of the occupied and virtual orbitals separately. In order to converge the HF reference determinant of LiH- H_2O , we use a Fock matrix mixing of 35 % in place of the default value (30% [97]) used for the other systems. The Brillouin zone shrinking factor determining the density in reciprocal space is set to 3 for the LiH- H_2O system. This is not as close to the thermodynamic limit as in Tsatsoulis et al. [96], but sufficient for our purpose. The remaining shrinking factors were 9 for neon and 8 for $2\text{H}_2\text{O}$ and ethylene.

The 6-31G basis set [98] obtained from the basis set exchange database [99,100] is used for the 1D ethylene system. For the hydrogen atoms in the $2\text{H}_2\text{O}$ and LiH- H_2O systems, we use a valence triple-zeta basis set with

polarisation functions [101], and for Li we use a 6-1G [102] basis set. The 6-31G basis set for oxygen [103] is used both in $2\text{H}_2\text{O}$ and LiH- H_2O .

The cc-pVDZ RI fitting basis by Weigend et al. [91] is used for ethylene and LiH- H_2O , while the cc-pVTZ RI fitting basis [91] is used for neon and $2\text{H}_2\text{O}$, both for XDEC and CRYSCOR calculations. In order to avoid linear dependence, exponents less than 0.4 are removed for LiH- H_2O and less than 0.1 for the remaining systems.

We use the frozen-core approximation in the correlation treatment for all cases except 3D neon. The MP2 equations are solved iteratively using fixed point iterations until the residual norm is below a numerical threshold set to 10^{-10} , except for the $2\text{H}_2\text{O}$ system where it is determined dynamically from the FOT as $10^{-3} \times \text{FOT}$. In the expansive step of the fragment optimisation, we include by default a minimum of 6 new orbitals in the local subspace in each iteration. For 3D neon, however, we include a minimum of 10 new orbitals in order to be consistent with Ref. [44].

5. Results and discussion

The XDEC-RI implementation differs fundamentally from our previous work [44], so we first validate the implementation by comparing fragment and pair-fragment MP2 energies for 3D neon, ethylene, and $2\text{H}_2\text{O}$ with results obtained with CRYSCOR [42,51,57,68–71] and the original XDEC implementation [44]. In all cases, the occupied space is completely fragmented, i.e. each occupied orbital defines a fragment. The XDEC-RI truncation parameters are set to $\xi_0 = 10^{-14}$ and $\xi_1 = 10^{-14}$, and the attenuation is $\omega = 0.1 \text{ Bohr}^{-1}$. For 3D neon, we use the XDEC-RI approach for the virtual space, converging each fragment energy to $\text{FOT} = 10^{-4}$ Hartree, while a CIM-like fragmentation is used for ethylene and $2\text{H}_2\text{O}$, where for each occupied orbital the amplitudes are solved within $d_{\text{occ}} = d_{\text{virt}} = 20 \text{ Bohr}$ and thereafter cast into pairwise energy contributions. The results, depicted as functions of the pair separation up to 20 Bohr in Figure 4, show excellent agreement across the implementations.

To gain more insight into the impact of the attenuated RI approximation, we present total MP2 energies for 1D neon in Table 1 for various values of the attenuation parameter ω , using the CIM result from

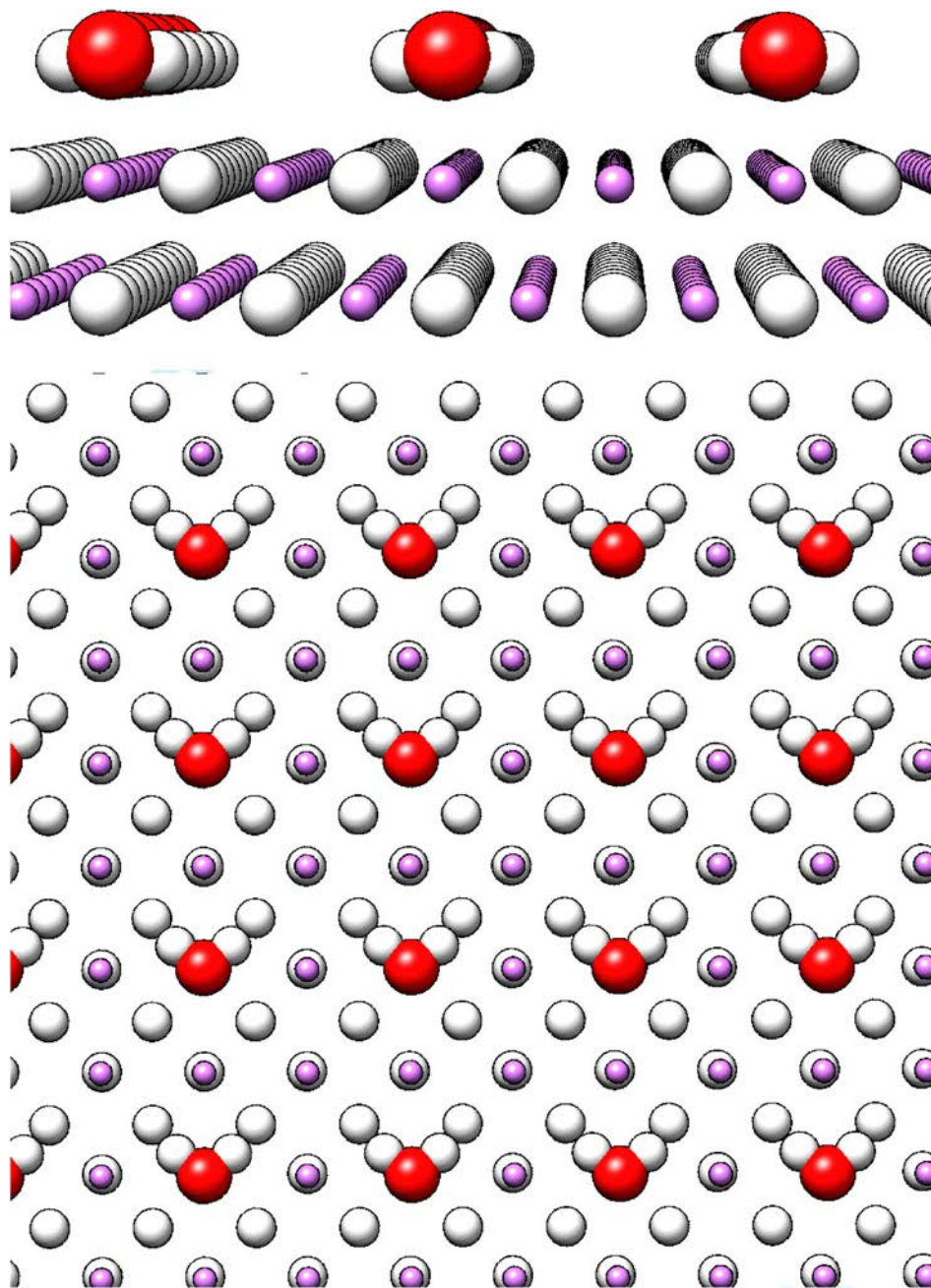


Figure 3. Geometrical setup of the LiH-H₂O adsorption system. A sideview is shown on top, while a top-down view is shown below. The equilibrium geometry was taken from the work of Tsatsoulis et al. [96]

Ref. [43] as reference. With a cc-pVDZ fitting basis [91], XDEC-parameters of $\xi_0 = 10^{-12}$, $\xi_1 = 10^{-12}$, FOT = 10^{-6} Hartree and orbital increment 10 [44], we find fairly good results already at $\omega = 100.0 \text{ Bohr}^{-1}$. This is likely due to the fact that most of the significant correlation effects in this system occur internally on the atoms. The results are numerically the same as CIM for $\omega = 0.5 \text{ Bohr}^{-1}$, and we observe no significant change below this value for this system.

Ideally, we would like the effect of the attenuation to be minimal in comparison to the ones caused by the

distance cutoffs, so a closer scrutiny of the attenuation dependence in the energy is warranted. We therefore compute the CIM-like MP2 energies for ethylene and 3D neon for a range of ω -values, as shown in Figure 5. Again, we obtain validating results from CRYSCOR using the same auxiliary basis with a Coulomb-metric fit, PAOs for the virtual orbital space and a local excitation domain for all strong pairs which includes all virtual orbitals associated with the 25 nearest neighbouring atoms for ethylene, and 10 nearest neighbours for neon.

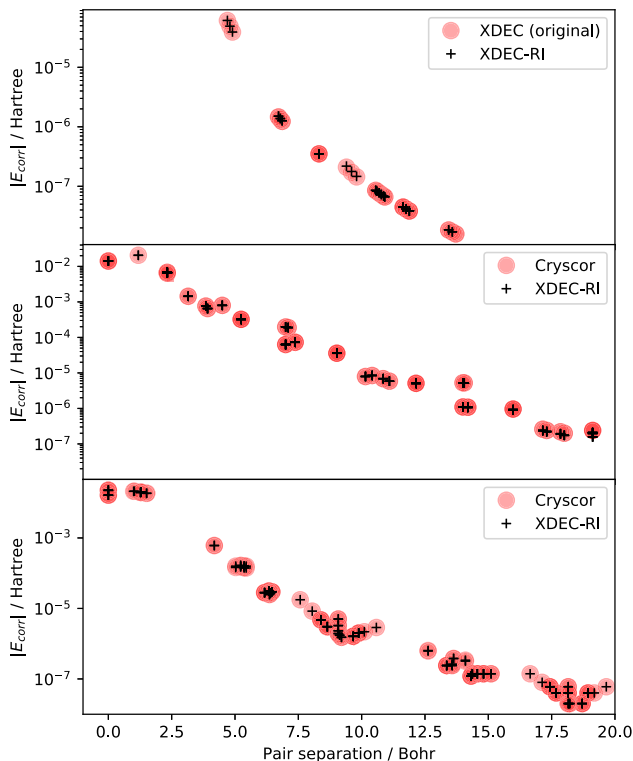


Figure 4. Validation of pair-fragment energies for (top to bottom) 3D neon, ethylene and $2\text{H}_2\text{O}$ by comparison to CRYSCOR and the original XDEC implementation. The fragment energy for 3D neon is not included.

With a cutoff for the local subspace $d_{\text{virt}} = 15.0$ Bohr for ethylene and $d_{\text{virt}} = 6.0$ Bohr for 3D neon, and a range of d_{occ} -values, we find that both cases show the energy approaching in a smooth fashion the Coulomb-metric fit as indicated by the CRYSCOR results when ω is decreased. The dependence on the attenuation parameter is similar to that reported for molecules in the past [50,104], with a stable region for $\omega \leq 0.3 \text{ Bohr}^{-1}$. In the region from 0.3 Bohr^{-1} to 10.0 Bohr^{-1} , the energy changes smoothly before it stabilises not too far from the values obtained close to the Coulomb-metric fit.

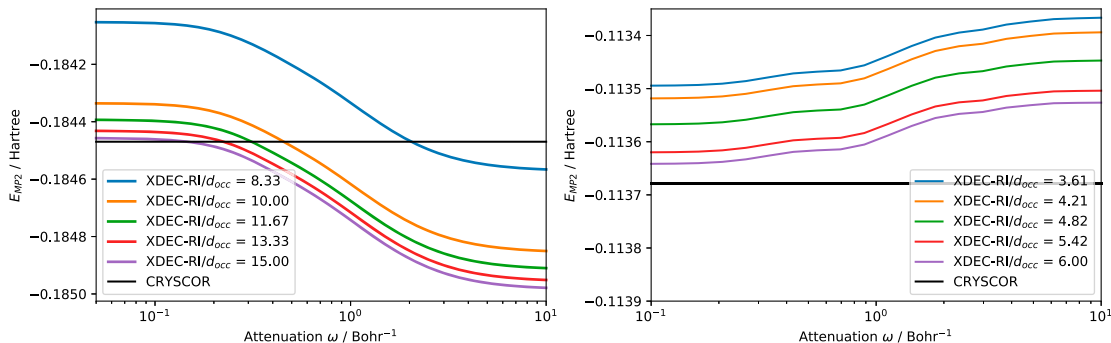


Figure 5. Dependence of the LMP2 energy on the attenuation parameter ω for ethylene (left) and 3D neon (right). For ethylene, we had $d_{\text{virt}} = 15.0$ Bohr, while for 3D neon we had $d_{\text{virt}} = 6.0$ Bohr.

Table 1. Total MP2 energy for 1D neon compared to CIM results from Wang et al. [43].

	Attenuation	$E_{\text{MP2}} / \text{Hartree}$
XDEC-RI	100.0	-0.114314
	10.0	-0.114319
	1.0	-0.114351
	0.5	-0.114363
	0.1	-0.114363
	0.07	-0.114363
CIM	0.05	-0.114363
	-	-0.114363

Variations in the cohesion energy for ethylene following changes in attenuation in the region below $\omega = 0.3 \text{ Bohr}^{-1}$ is found to be less than 1% relative to $796.5 \mu\text{Ha}$, obtained at $\omega = 0.05 \text{ Bohr}^{-1}$. In general, the changes in energy due to attenuation below 0.3 Bohr^{-1} are fairly small in comparison to the changes in the cutoff parameter d_{occ} , so we conclude that we may safely choose $\omega \leq 0.3 \text{ Bohr}^{-1}$ in our further examination.

The features of the curves in Figure 5 corresponding to the various distance cutoffs appear more or less internally unchanged, while they are simply shifted in energy as the distance cutoff changes. This indicates that the correlation energy distributes similarly across the pairs independent of the attenuation parameter, but it does not decisively tell us whether or not the attenuation could cause artefacts in the PES beyond the ones we expect from the fragmentation. Since the position of each orbital remains fixed after the localisation, we expect abrupt changes in energy to occur at the same distance cutoffs regardless of the various other parameters used in the approximation. Thus, in order to rule out such attenuation-related effects in the energy, we compute for $\omega = 0.1 \text{ Bohr}^{-1}$ and $\omega = 0.2 \text{ Bohr}^{-1}$, both safely within the limit of $\omega = 0.3 \text{ Bohr}^{-1}$, the CIM-like energies for $2\text{H}_2\text{O}$ for a range of cutoffs. This is a slightly more correlated system, since it features hydrogen bonds both within and between the cells. In the results, presented in Figure 6, we find our first indications of abrupt

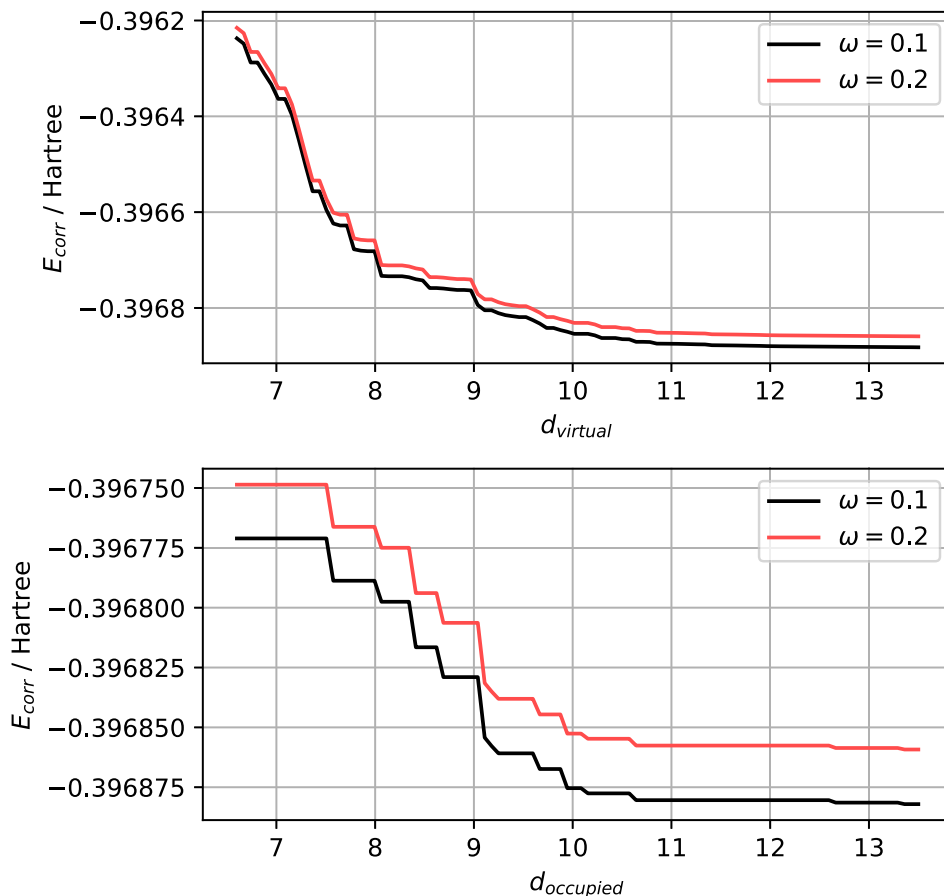


Figure 6. The XDEC-RI MP2 energy of $2\text{H}_2\text{O}$ for various distance cutoffs. In the upper panel, $d_{occ} = 13$ Bohr and in the lower panel $d_{virt} = 13$ Bohr.

energetic changes, represented by an irregular convergence pattern featuring cliffs and plateaus as the cutoff parameters are increased in both the virtual and occupied direction. The cliffs in this system are an order of magnitude larger in the virtual direction, on the order of 10^{-4} Hartree, as compared to 10^{-5} Hartree in the occupied direction. Comparing the curves for each ω , we find that discontinuities along both d_{occ} and d_{virt} remain stationary when the attenuation parameter is changed. This suggests that any discontinuities we see here are independent of the attenuation parameter for $\omega \leq 0.3 \text{ Bohr}^{-1}$, and are primarily caused by discrete changes in the local subspace.

A more complete picture of the convergence pattern can be revealed by varying d_{occ} and d_{virt} simultaneously. In order to illustrate the theoretical considerations we have made in regards to discontinuities, we therefore use our code to compute energies for a range of distance cutoff parameters for ethylene and $2\text{H}_2\text{O}$, shown in Figure 7. In this way, the local MP2 energy can be shown as a surface, approaching the exact result as d_{occ} and d_{virt} becomes large. Both systems again show the expected non-smooth yet monotonically decreasing energy with

respect to the domain sizes in the form of cliffs and plateaus across the surface. The abrupt changes occur at the same distances for one of the cutoff parameters seemingly independently of the other. This feature suggests that it is possible to converge the energy first with regards to one cutoff, thereafter the other, in contrast to repeated successive convergence of these two parameters conventionally used in DEC methods [19,44].

We have indicated the *mean* number of occupied or virtual orbitals per local subspace that are included in the calculation for some chosen cutoffs, in order to highlight that the sharp cliffs do not necessarily correspond to distances at which a large number of orbitals enter the calculation. Rather, it shows there are certain orbitals that yield more significant contributions to the local correlation energy, whether it be through direct contributions or indirect buffer effects. These are scattered throughout the neighbourhood of the reference cell, not necessarily ordered by distance.

Large contributions to the energy from close pairs are clearly present in both cases, while the contributions in the virtual direction tend to be more evenly spread out over larger intervals. This latter effect can likely be

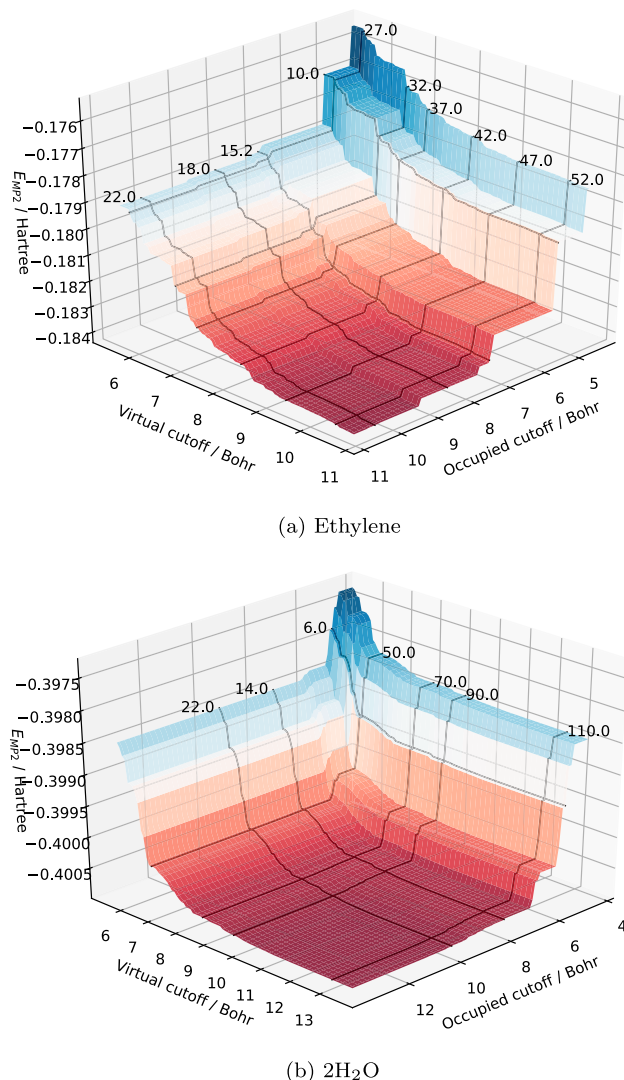


Figure 7. Convergence of the CIM-like energy of the ethylene system (top) and a 2H₂O system with respect to occupied and virtual cutoff distances. The numbered lines on the surface indicate the mean number of virtual and occupied orbitals per occupied that are included per local subspace at a given truncation. Note that there are energies from several local subspaces superimposed in these figures, thus the number of orbitals may be fractional. (a) Ethylene. (b) 2H₂O.

explained by the centres of the virtual orbitals being more arbitrarily dispersed in comparison to the occupied orbitals who will likely be positioned close to atoms or bonding sites in the lattice. Within the DEC formalism, it is customary to assign orbitals to atomic sites by means of for instance Mulliken or Löwdin charges [18,21,44]. If we instead assign the virtual orbitals to their closest atomic centres and include all virtual orbitals for all atoms inside d_{virt} in the local subspace, we obtain for ethylene the results shown in Figure 8 in support of this explanation. Here, we see clearly that the energy for the atomic association features more distinct plateaus and cliffs in

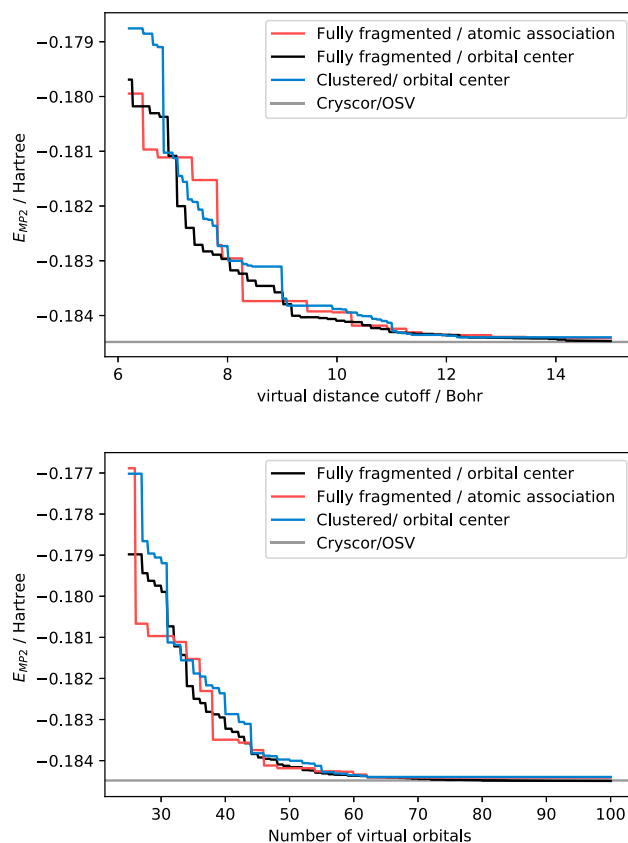


Figure 8. The convergence of the virtual space for ethylene; a comparison of a setup where orbitals are associated with atoms and the default orbital-centred approach. We have also included results where the occupied space was partitioned into two groups based on their nearest neighbour rather than into individual orbitals. The occupied cutoff is fixed at $d_{\text{occ}} = 15.0$ Bohr. As a reference, Cryscor results for OSV tolerance 10^{-5} are shown. The average size of the OSV orbital domains is 26.

the virtual direction as compared to the one where the orbital centres are used directly. This effect persists independently of whether we count distance or number of virtual orbitals along the x -axis, which further reinforces the point that the significance of each orbital with respect to the correlation energy is not simply a function of the distance.

We see similar irregularities emerge if we allow for more than one orbital per fragment or cluster, as shown for ethylene and 2H₂O in Figures 8 and 9, respectively. For ethylene, we here compare the outcome of a CIM-like MP2 calculation where the occupied space has been subdivided into two clusters based on distance as outlined in the implementational details, to the completely fragmented calculation. Similarly, for water, we compare two clusters to the complete fragmentation. In both cases, we find slightly sharper cliffs and flatter plateaus for the clustered approach. The effect is independent of whether

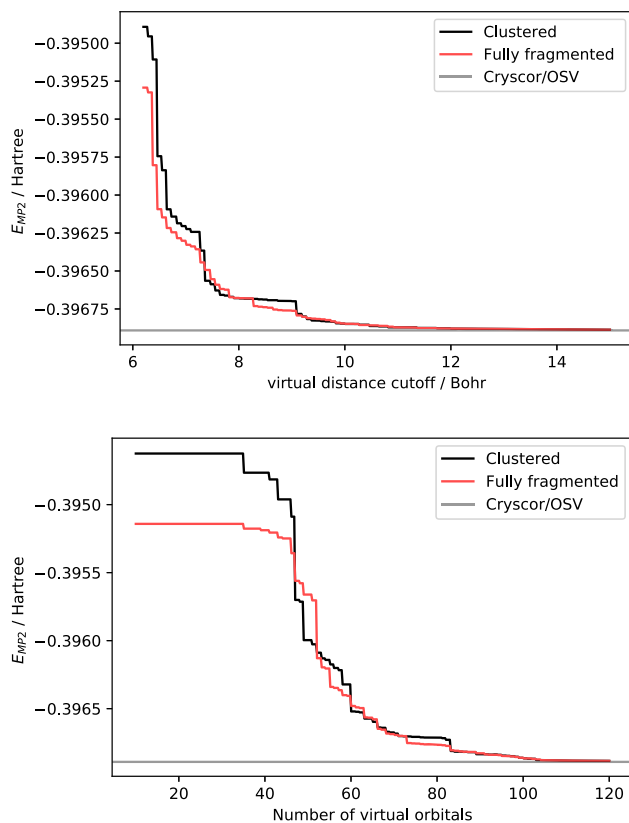


Figure 9. The convergence of the virtual space for $2\text{H}_2\text{O}$; a comparison between a setup where the occupied space is fully fragmented into individual orbitals and another setup where they are partitioned into 4 groups based on their nearest neighbours. The occupied cutoff is fixed at $d_{\text{occ}} = 15.0$ Bohr. As a reference, *Cryscor* results for OSV tolerance 10^{-5} are shown.

distance or number of virtual orbitals is used for the x -axis. When spatially non-coincident occupied orbitals are grouped into the same fragment or cluster, a purely distance-based ordering of the virtual space can result in a more irregular convergence pattern than that of a totally fragmented occupied space, likely due to the fact that being close does not necessarily guarantee strong coupling to the same virtual subspace.

These results show that the discontinuities are clearly dependent on the fragmentation, where both the clustering of occupied orbitals and the common practice of associating the virtual orbitals to atoms may increase the likelihood of false convergence of the MP2 energy with respect to d_{virt} .

We then focus on geometry dependence in the PES by considering the adsorption of a water molecule on a surface of LiH, as shown in Figure 3. We keep the internal geometry of H_2O fixed, while its distance perpendicular to the surface is varied between -1 to 2 Bohr relative to its equilibrium. The equilibrium geometry is taken from Ref. [96], where the water molecule has been relaxed at the LiH surface at the KS-DFT level with the

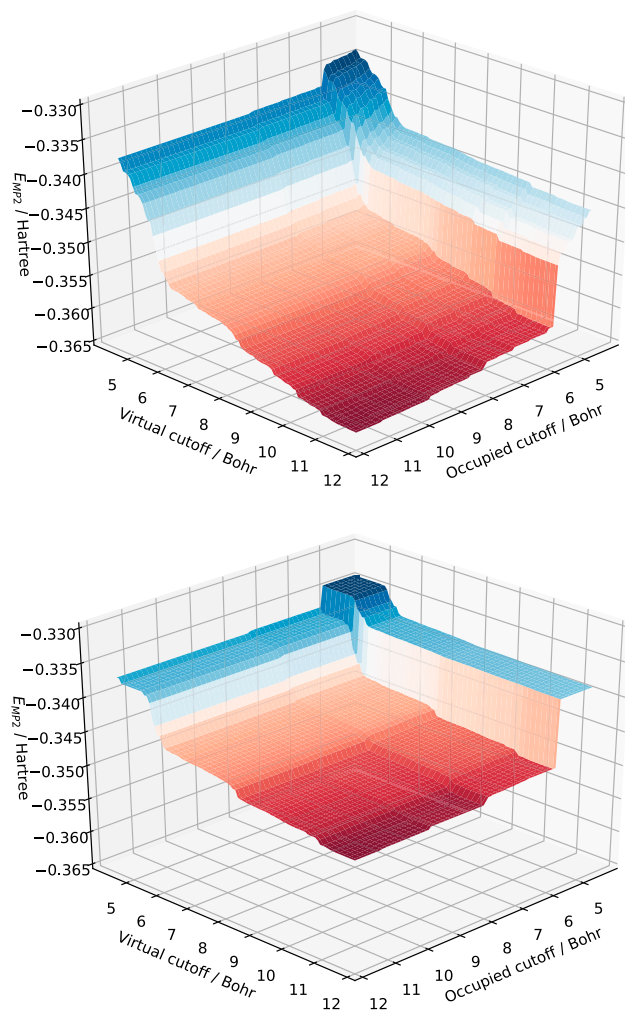


Figure 10. The MP2 energy as a function of the local subspace definition for a displacement of H_2O from equilibrium perpendicular to the surface of $\Delta z_{\text{H}_2\text{O}} = -1.115$ Bohr (above) and $\Delta z_{\text{H}_2\text{O}} = 1.871$ Bohr (below). As the water molecule approach the surface, the energy decreases and the sharp discontinuities becomes smeared out – or eroded – resulting in a more smooth convergence pattern.

Perdew–Burke–Ernzerhof (PBE) XC functional [105]. We let $\omega = 0.15$ Bohr $^{-1}$.

As expected, varying the geometry results in significant changes in the convergence pattern, as shown in Figure 10. The position of the orbitals in the LiH surface moves as the water molecule approaches the surface, resulting in a type of erosion of the sharp cliffs and plateaus. Also, the MP2 energy is generally decreased when water comes close as the interactions with the surface intensifies. In Figure 11, we show the effect of various choices for distance truncations of the local subspace.

Discontinuities are visible even for the largest subspace, and upon inspection, they are found to be on the order of 10^{-4} Hartree. More discontinuities emerge for

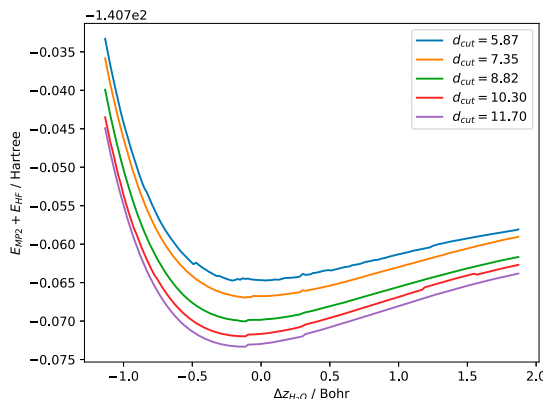
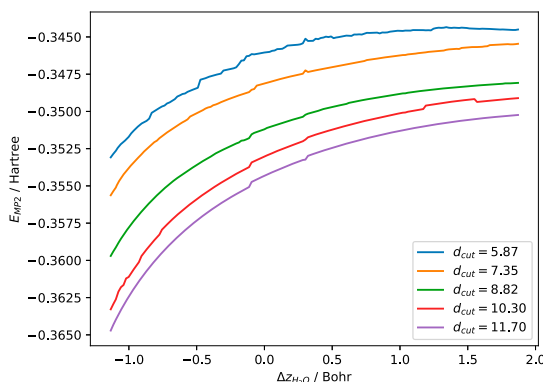


Figure 11. MP2 energy (left) and MP2 + reference energy (right) for the LiH-H₂O surface. The x-axis shows the displacement of the water molecule relative to the relaxed geometry perpendicular to the surface. The distance truncation of the local subspace is in this case $d_{\text{cut}} = d_{\text{occ}} = d_{\text{virt}}$.

the smaller domains. We can observe at least two discontinuities at approximately -0.1 and 0.3 Bohr which persists across the various cutoffs. The one at -0.1 Bohr is positioned awkwardly close to the minimum of the total energy, highlighting exactly how these kinds of effects can influence the results. In this case, the minimum of the potential curve is shifted slightly towards negative $\Delta z_{\text{H}_2\text{O}}$ due to the presence of a discontinuity.

The MP2 adsorption energy is computed using a similar definition as in Refs [96], where only inter H₂O-surface pairs up to our maximum cutoff threshold are included in the energy. In terms of the energy expression in Equation (9), this corresponds to

$$E_{\text{ads,MP2}} = 2 \sum_{i \in \Omega_0} \sum_{j \in \mathcal{O} \setminus \Omega_0} \sum_{ab \in \mathcal{V}} E_{ij}^{ab}, \quad (45)$$

where Ω_0 signifies the set of occupied orbitals situated at the H₂O molecule in the reference cell.

From the convergence pattern of the MP2 adsorption energy at equilibrium ($\Delta z_{\text{H}_2\text{O}} = 0$ Bohr), shown in Figure 12, we can clearly see a step-like pattern for pairs at distances below 8 Bohr where significant contributions to the energy are present. These significant contributions can be attributed to surface-molecule pairs close to the water molecule, as to be expected with localised orbitals. An unexpected effect is, however, present in the virtual cutoff, where sharp cliffs are present beyond distances $d_{\text{virt}} \geq 8$ Bohr.

The effect on the convergence pattern from the moving water molecule is shown in Figure 13. Also here we can identify the step-like pattern for close pairs, and the more distant contributions from the virtual space. The movement of the water molecule reveals a distinction between these effects; while the discontinuities in the occupied direction move along with the geometry, the virtual discontinuities appear to remain stationary. The

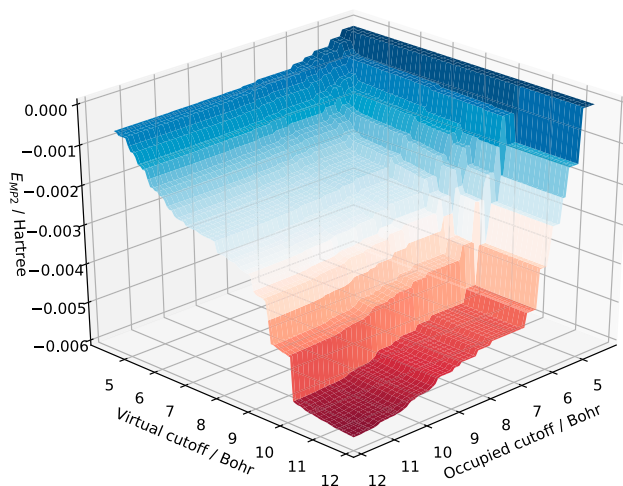


Figure 12. MP2 adsorption energy in Hartrees (LiH-H₂O surface-molecule pairs within cutoff distances specified) as a function of the distance truncation parameters at equilibrium ($\Delta z_{\text{H}_2\text{O}} = 0$ Bohr).

distance at which they occur is just below the lattice parameter (10.91 Bohr) and suggests that these contributions may be attributed to virtual orbitals at neighbouring water molecules, or similarly virtual orbitals at the corresponding Li- or H-atoms in neighbouring cells.

When compared to Figure 12, we find that these distant contributions from the virtual space persist also for small occupied cutoffs, meaning that the distant virtual orbitals are significant even when both occupied orbitals are close to the origin. This could be an effect pertaining to the LVO representation of the virtual space, but further studies are required in order to determine whether or not it would be resolved by replacing them with PAOs or OSVs.

The DEC approach of converging the domains can in principle be applied to any of the fragmentation schemes under consideration. In order to demonstrate this, we

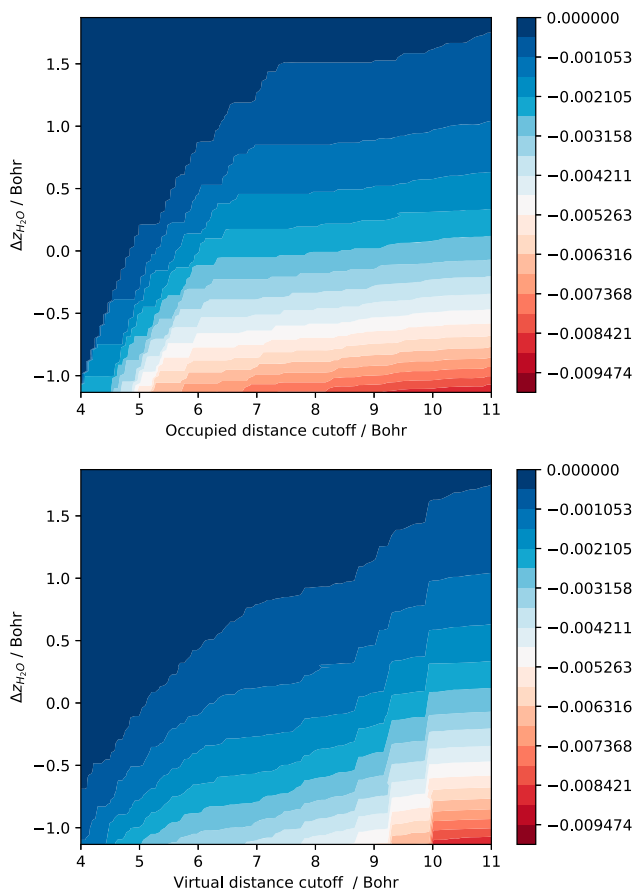


Figure 13. MP2 adsorption energy in Hartrees (LiH-H₂O surface – molecule pairs within cutoff distances specified) as a function of the distance truncation parameters in the occupied direction when d_{virt} is at its maximum value (top) and similarly as a function of d_{virt} when d_{occ} is at its maximum value (bottom). The y axis shows the vertical offset from equilibrium of the water molecule.

finally investigate the relationship between the magnitude of the discontinuities and the FOT in the 2H₂O-system by converging the CIM-like MP2 energy successively in the virtual and occupied directions as described in Ref. [44]. We increase the spaces by a minimum of 6 orbitals per expansive step. As the lattice parameter is altered, one H₂O molecule is moved such that the bond angles are preserved. The results are shown in Figure 14. We finally find that the magnitude of the discontinuities appears to be proportional to the FOT, which confirms that the discontinuities indeed can be systematically suppressed in this manner.

6. Concluding remarks

Discontinuities are inherently present in periodic PESs produced with local correlation methods. The overall source of these discontinuities is discrete changes in the orbital spaces, and as long as the spaces are chosen sufficiently large these problems may be insignificant. Exactly

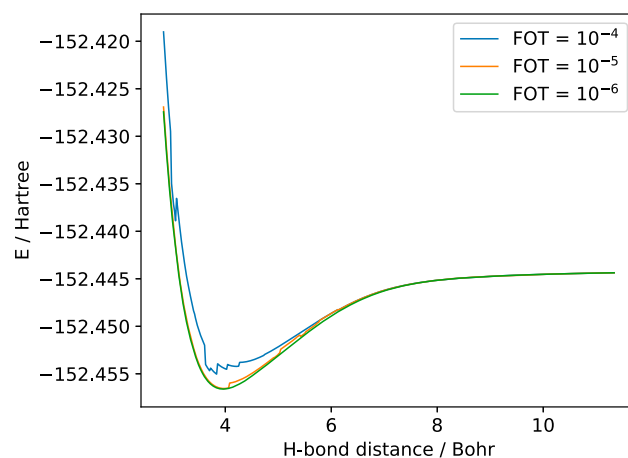


Figure 14. Adaptive CIM approach for a chain of H₂O (two molecules per cell). The figure demonstrates that discontinuities can be systematically suppressed for a CIM-based fragmentation scheme by converging the energy of each cluster analogous to what is done in XDEC.

what constitutes sufficiently large is, however, system dependent and difficult to determine in a black box manner. The DEC approach of converging the energies of each subspace can in principle provide control of the magnitude of the discontinuities for most fragmentation schemes, but the practical complications due to the irregular convergence behaviour represent a challenge. Ideally, one would like every new set of orbitals that enters the subspace to yield an energy contribution smaller than the preceding ones. Our results confirm as expected that this is not the situation for the purely distance-based selection of the WFs, but that in large enough increments the overall behaviour is convergent.

The convergence behaviour is dependent on several fragmentation-specific choices that are easy to control. Notably, the risk of false convergence can be reduced by choosing more fine-grained fragmentation and avoid coincidental position definitions. Furthermore, the balanced treatment of the virtual space of the Pulay–Sæbø pairs and DEC pair fragments is natural to include in periodic systems, for example by means of OSVs [57] or a global fitting scheme, and may further reduce the risk of discontinuities from the virtual cutoff.

In order to converge the local subspace, efficient handling of the integrals is crucial. Such efficiency can be achieved with a global fitting scheme as demonstrated in this work. While the initial partial contraction of the occupied space is computationally demanding, the subsequent contraction of virtual spaces and calculation of integrals is relatively cheap, and the intermediate contraction tensor contains all the required information to obtain all ERIs for the crystal.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

A. S. Hansen  <http://orcid.org/0000-0003-0962-3143>

T. B. Pedersen  <http://orcid.org/0000-0001-8967-6055>

References

- [1] W. Kohn and L.J. Sham, *Phys. Rev.* 140, A1133–A1138 (1965). doi:10.1103/PhysRev.140.A1133
- [2] C.C.J. Roothaan, *Rev. Mod. Phys.* 23, 69–89 (1951). doi:10.1103/RevModPhys.23.69
- [3] R.J. Bartlett and M. Musiał, *Rev. Mod. Phys.* 79, 291–352 (2007). doi:10.1103/RevModPhys.79.291
- [4] O. Sinanoğlu, *Adv. Chem. Phys.* 6, 315–412 (1964). doi:10.1002/9780470143520.ch7
- [5] R.K. Nesbet, *Adv. Chem. Phys.* 9, 321–363 (1965). doi:10.1002/9780470143551.ch4
- [6] J.M. Cullen and M.C. Zerner, *J. Chem. Phys.* 77, 4088–4109 (1982). doi:10.1063/1.444319
- [7] P. Pulay, *Chem. Phys. Lett.* 100, 151–154 (1983). doi:10.1016/0009-2614(83)80703-9
- [8] S. Sæbø and P. Pulay, *Chem. Phys. Lett.* 113, 13–18 (1985). doi:10.1016/0009-2614(85)85003-X
- [9] P. Pulay and S. Sæbø, *Theor. Chim. Acta* 69, 357–368 (1986). doi:10.1007/BF00526697
- [10] P. Pulay and S. Sæbø, *J. Chem. Phys.* 914, 914–922 (1987). doi:10.1063/1.452293
- [11] W. Meyer, *J. Chem. Phys.* 58, 1017–1035 (1973). doi:10.1063/1.1679283
- [12] F. Neese, F. Wennmohs and A. Hansen, *J. Chem. Phys.* 130, 114108 (2009). doi:10.1063/1.3086717
- [13] F. Neese, A. Hansen and D.G. Liakos, *J. Chem. Phys.* 131, 064103 (2009). doi:10.1063/1.3173827
- [14] C. Riplinger and F. Neese, *J. Chem. Phys.* 138, 034106 (2013). doi:10.1063/1.4773581
- [15] P. Pinski, C. Riplinger, E.F. Valeev and F. Neese, *J. Chem. Phys.* 143, 034108 (2015). doi:10.1063/1.4926879
- [16] C. Riplinger, P. Pinski, U. Becker, E.F. Valeev and F. Neese, *J. Chem. Phys.* 144, 024109 (2016). doi:10.1063/1.4939030
- [17] H.-J. Werner, G. Knizia, C. Krause, M. Schwilk and M. Dornbach, *J. Chem. Theory Comput.* 11, 484–507 (2015). doi:10.1021/ct500725e
- [18] M. Ziolkowski, B. Jansik, T. Kjærgaard and P. Jørgensen, *J. Chem. Phys.* 133, 014107 (2010). doi:10.1063/1.3456-535
- [19] K. Kristensen, M. Ziolkowski, B. Jansik, T. Kjærgaard and P. Jørgensen, *J. Chem. Theory Comput.* 7, 1677–1694 (2011). doi:10.1021/ct200114k
- [20] I.-M. Høyvik, K. Kristensen, B. Jansik and P. Jørgensen, *J. Chem. Phys.* 136, 014105 (2012). doi:10.1063/1.3667266
- [21] K. Kristensen, P. Jørgensen, B. Jansik, T. Kjærgaard and S. Reine, *J. Chem. Phys.* 137, 114102 (2012). doi:10.1063/1.4752432
- [22] K. Kristensen, T. Kjærgaard, I.-M. Høyvik, P. Etenhuber, P. Jørgensen, B. Jansik, S. Reine and J. Jakowski, *Mol. Phys.* 111, 1196–1210 (2013). doi:10.1080/00268976.2013.783941
- [23] S. Li, J. Ma and Y. Jiang, *J. Comput. Chem.* 23 (2), 237–244 (2002). doi:10.1002/(ISSN)1096-987X
- [24] S. Li, J. Shen, W. Li and Y. Jiang, *J. Chem. Phys.* 125 (7), 074109 (2006). doi:10.1063/1.2244566
- [25] H.-J. Werner and K. Pflüger, *Annu. Rep. Comput. Chem.* 2, 53–80 (2006). doi:10.1016/S1574-1400(06)02004-4
- [26] F. Bloch, *Z. Phys.* 52, 555–600 (July 1929). doi:10.1007/BF01339455
- [27] R.D.L. Kronig, W.G. Penney and R.H. Fowler, *Proc. Math. Phys. Eng. Sci.* 130, 499–513 (1931). doi:10.1098/rspa.1931.0019
- [28] N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Saunders, Holt-Saunders, 1976).
- [29] A. Grüneis, M. Marsman, J. Harl, L. Schimka and G. Kresse, *J. Chem. Phys.* 131, 154115 (2009). doi:10.1063/1.3250347
- [30] G.H. Booth, A. Grüneis, G. Kresse and A. Alavi, *Nature* 493, 365–370 (2013). doi:10.1038/nature11770
- [31] T.D. Kühne, M. Iannuzzi, M. Del Ben, V.V. Rybkin, P. Seewald, F. Stein, T. Laino, R.Z. Khaliullin, O. Schütt and F. Schiffmann, *J. Chem. Phys.* 152, 194103 (2020). doi:10.1063/5.0007045
- [32] J.-Q. Sun and R.J. Bartlett, *J. Chem. Phys.* 104, 8553–8565 (1996). doi:10.1063/1.471545
- [33] S. Hirata and S. Iwata, *J. Chem. Phys.* 109, 4147–4155 (1998). doi:10.1063/1.477020
- [34] S. Hirata and R.J. Bartlett, *J. Chem. Phys.* 112, 7339–7344 (2000). doi:10.1063/1.481372
- [35] P.Y. Ayala, K.N. Kudin and G.E. Scuseria, *J. Chem. Phys.* 115, 9698–9707 (2001). doi:10.1063/1.1414369
- [36] J. McClain, Q. Sun, G.K.-L. Chan and T.C. Berkelbach, *J. Chem. Theory Comput.* 13, 1209–1218 (2017). doi:10.1021/acs.jctc.7b00049
- [37] G.H. Wannier, *Phys. Rev.* 52, 191 (1937). doi:10.1103/PhysRev.52.191
- [38] W. Kohn, *Phys. Rev.* 115, 809–821 (1959). doi:10.1103/PhysRev.115.809
- [39] D. Fiorenza, D. Monaco and G. Panati, *Ann. Henri Poincaré* 17, 63–97 (2016). doi:10.1007/s00023-015-0400-6
- [40] N. Marzari, A.A. Mostofi, J.R. Yates, I. Souza and D. Vanderbilt, *Rev. Mod. Phys.* 84, 1419–1475 (2012). doi:10.1103/RevModPhys.84.1419
- [41] C.M. Zicovich-Wilson, R. Dovesi and V.R. Saunders, *J. Chem. Phys.* 115, 9708–9719 (2001). doi:10.1063/1.141-5745
- [42] C. Pisani, L. Maschio, S. Casassa, M. Halo, M. Schütz and D. Usvyat, *J. Comput. Chem.* 29, 2113–2124 (2008). doi:10.1002/jcc.v29:13
- [43] Y. Wang, Z. Ni, W. Li and S. Li, *J. Chem. Theory Comput.* 15, 2933–2943 (2019). doi:10.1021/acs.jctc.8b01200
- [44] E. Rebolini, G. Baardsen, A.S. Hansen, K.R. Leikanger and T.B. Pedersen, *J. Chem. Theory Comput.* 14, 2427–2438 (2018). doi:10.1021/acs.jctc.8b00021
- [45] A.S. Hansen, G. Baardsen, E. Rebolini, L. Maschio and T.B. Pedersen, *Mol. Phys.* 1–11 (2020). doi:10.1080/00268976.2020.1733118
- [46] M. Lorenz, D. Usvyat and M. Schütz, *J. Chem. Phys.* 134, 094101 (2011). doi:10.1063/1.3554209

- [47] M. Lorenz, L. Maschio, M. Schütz and D. Usvyat, *J. Chem. Phys.* 137, 204119 (2012). doi:10.1063/1.4767775
- [48] M. Milko, J. Noga and Š. Varga, *Int. J. Quantum Chem.* 107, 2158–2168 (2007). doi:10.1002/(ISSN)1097-461X
- [49] L. Maschio, *J. Chem. Theory Comput.* 7, 2818–2830 (2011). doi:10.1021/ct200352g
- [50] Y. Jung, A. Sodt, P.M. Gill and M. Head-Gordon, *Proc. Natl. Acad. Sci. U.S.A.* 102, 6692–6697 (2005). doi:10.1073/pnas.0408475102
- [51] L. Maschio, D. Usvyat, F.R. Manby, S. Casassa, C. Pisani and M. Schütz, *Phys. Rev. B* 76, 075101 (2007). doi:10.1103/PhysRevB.76.075101
- [52] D. Usvyat, L. Maschio, F.R. Manby, S. Casassa, M. Schütz and C. Pisani, *Phys. Rev. B* 76, 075102 (2007). doi:10.1103/PhysRevB.76.075102
- [53] F.R. Manby, editors, *Accurate Condensed-Phase Quantum Chemistry* (Boca Raton, CRC Press, 2010).
- [54] N.J. Russ and T.D. Crawford, *J. Chem. Phys.* 121, 691–696 (2004). doi:10.1063/1.1759322
- [55] R.A. Mata and H.-J. Werner, *J. Chem. Phys.* 125, 184110 (2006). doi:10.1063/1.2364487
- [56] J.E. Subotnik, A. Sodt and M. Head-Gordon, *J. Chem. Phys.* 128, 034103 (2008). doi:10.1063/1.2821124
- [57] D. Usvyat, L. Maschio and M. Schütz, *J. Chem. Phys.* 143, 102805 (2015). doi:10.1063/1.4921301
- [58] M.S. Lee, P.E. Maslen and M. Head-Gordon, *J. Chem. Phys.* 112, 3592–3601 (2000). doi:10.1063/1.480512
- [59] P. Löwdin, *J. Chem. Phys.* 18, 365–375 (1950). doi:10.1063/1.1747632
- [60] G. Hetzer, P. Pulay and H.-J. Werner, *Chem. Phys. Lett.* 290, 143–149 (1998). doi:10.1016/S0009-2614(98)00491-6
- [61] H.-J. Werner, *J. Chem. Phys.* 145, 201101 (2016). doi:10.1063/1.4968595
- [62] J. Yang, Y. Kurashige, F.R. Manby and G.K. Chan, *J. Chem. Phys.* 134, 044123 (2011). doi:10.1063/1.3528935
- [63] Y. Kurashige, J. Yang, G.K.-L. Chan and F.R. Manby, *J. Chem. Phys.* 136, 124106 (2012). doi:10.1063/1.3696962
- [64] D. Kats and F.R. Manby, *J. Chem. Phys.* 138, 144101 (2013). doi:10.1063/1.4798940
- [65] C. Krause and H.-J. Werner, *Phys. Chem. Chem. Phys.* 14, 7591–7604 (2012). doi:10.1039/c2cp40231a
- [66] S. Saebø and P. Pulay, *Annu. Rev. Phys. Chem.* 44, 213 (1993). doi:10.1146/annurev.pc.44.100193.001241
- [67] H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby and M. Schütz, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2, 242–253 (2012). doi:10.1002/wcms.82
- [68] C. Pisani, M. Busso, G. Capecchi, S. Casassa, R. Dovesi, L. Maschio, C. Zicovich-Wilson and M. Schütz, *J. Chem. Phys.* 122, 094113 (2005). doi:10.1063/1.1857479
- [69] M. Halo, S. Casassa, L. Maschio and C. Pisani, *Phys. Chem. Chem. Phys.* 11, 586–592 (2009). doi:10.1039/B812870G
- [70] C. Pisani, M. Schütz, S. Casassa, D. Usvyat, L. Maschio, M. Lorenz and A. Erba, *Phys. Chem. Chem. Phys.* 14, 7615–7628 (2012). doi:10.1039/c2cp23927b
- [71] R. Dovesi, A. Erba, R. Orlando, C.M. Zicovich-Wilson, B. Civalleri, L. Maschio, M. Rérat, S. Casassa, J. Baima and S. Salustro, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 8, e1360 (2018). doi:10.1002/wcms.2018.8.issue-4
- [72] H. Stoll, *Phys. Rev. B* 46, 6700 (1992). doi:10.1103/PhysRevB.46.6700
- [73] B. Paulus, *Phys. Rep.* 428, 1–52 (2006). doi:10.1016/j.physrep.2006.01.003
- [74] M. Alessio, D. Usvyat and J. Sauer, *J. Chem. Theory Comput.* 15, 1329–1344 (2018). doi:10.1021/acs.jctc.8b01122
- [75] O. Masur, M. Schütz, L. Maschio and D. Usvyat, *J. Chem. Theory Comput.* 12, 5145–5156 (2016). doi:10.1021/acs.jctc.6b00651
- [76] H.-H. Lin, L. Maschio, D. Kats, D. Usvyat and T. Heine, *J. Chem. Theory Comput.* 16, 7100–7108 (2020). doi:10.1021/acs.jctc.0c00576
- [77] E.F. Valeev and J.T. Fermann, *Libint: A library for the evaluation of molecular integrals of many-body operators over Gaussian functions* (2017). <https://github.com/evaleev/libint>, Accessed: 2020-01-15.
- [78] C.V. van der Mee, S. Seatzu and G. Rodriguez, *Linear Algebra Appl.* 343, 355–380 (2002). doi:10.1016/S0024-3795(01)00441-4
- [79] H.J. Nussbaumer, *Fast Fourier Transform and Convolution Algorithms* (Springer, 1981)
- [80] C.L. Lawson, R.J. Hanson, D.R. Kincaid and F.T. Krogh, *ACM Trans. Math. Softw.* 5, 308–323 (1979). doi:10.1145/355841.355847
- [81] J.L. Whitten, *J. Chem. Phys.* 58, 4496–4501 (1973). doi:10.1063/1.1679012
- [82] E.J. Baerends, D. Ellis and P. Ros, *Chem. Phys.* 2, 41–51 (1973). doi:10.1016/0301-0104(73)80059-X
- [83] B.I. Dunlap, J. Connolly and J. Sabin, *J. Chem. Phys.* 71, 3396–3402 (1979). doi:10.1063/1.438728
- [84] B. Dunlap, *J. Chem. Phys.* 78, 3140–3142 (1983). doi:10.1063/1.445228
- [85] M. Feyereisen, G. Fitzgerald and A. Komornicki, *Chem. Phys. Lett.* 208, 359–363 (1993). doi:10.1016/0009-2614(93)87156-W
- [86] O. Vahtras, J. Almlöf and M. Feyereisen, *Chem. Phys. Lett.* 213, 514–518 (1993). doi:10.1016/0009-2614(93)89151-7
- [87] K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chem. Phys. Lett.* 240, 283–290 (1995). doi:10.1016/0009-2614(95)00621-A
- [88] K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theor. Chem. Acc.* 97, 119–124 (1997). doi:10.1007/s002140050244
- [89] F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.* 294, 143–152 (1998). doi:10.1016/S0009-2614(98)00862-8
- [90] C.-K. Skylaris, L. Gagliardi, N.C. Handy, A.G. Ioannou, S. Spencer and A. Willetts, *J. Mol. Struct. THEOCHEM* 501, 229–239 (2000). doi:10.1016/S0166-1280(99)00434-0
- [91] F. Weigend, A. Köhn and C. Hättig, *J. Chem. Phys.* 116, 3175–3183 (2002). doi:10.1063/1.1445115
- [92] J.P. Dombroski, S.W. Taylor and P.M. Gill, *J. Phys. Chem.* 100, 6272–6276 (1996). doi:10.1021/jp952841b
- [93] B. Dunlap, *J. Mol. Struct. THEOCHEM* 529, 37–40 (2000). doi:10.1016/S0166-1280(00)00528-5
- [94] A.M. Burow, M. Sierka and F. Mohamed, *J. Chem. Phys.* 131, 214101 (2009). doi:10.1063/1.3267858
- [95] A. Hoy and P.R. Bunker, *J. Mol. Spectrosc.* 74, 1–8 (1979). doi:10.1016/0022-2852(79)90019-5

- [96] T. Tsatsoulis, F. Hummel, D. Usvyat, M. Schütz, G.H. Booth, S.S. Binnie, M.J. Gillan, D. Alfè, A. Michaelides and A. Grüneis, *J. Chem. Phys.* 146, 204108 (2017). doi:10.1063/1.4984048
- [97] R. Dovesi, V.R. Saunders, C. Roetti, R. Orlando, C.M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N.M. Harrison, I.J. Bush, P. D'Arco, M. Llunell, M. Causá, Y. Noël, L. Maschio, A. Erba, M. Rerat and S. Casassa, *Crystal 17 user's manual* (2018). <https://www.crystal.unito.it/documentation.php>, Accessed: 2020-12-04.
- [98] W.J. Hehre, R. Ditchfield and J.A. Pople, *J. Chem. Phys.* 56, 2257–2261 (1972). doi:10.1063/1.1677527
- [99] D. Feller, *J. Comput. Chem.* 17, 1571–1586 (1996). doi:10.1002/(ISSN)1096-987X
- [100] K.L. Schuchardt, B.T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li and T.L. Windus, *J. Chem. Inf. Mod.* 47, 1045–1052 (2007). doi:10.1021/ci600510j
- [101] M.F. Peintinger, D.V. Oliveira and T. Bredow, *J. Comput. Chem.* 34, 451–459 (2013). doi:10.1002/jcc.23153
- [102] B. Civalleri, A.M. Ferrari, M. Llunell, R. Orlando, M. Mérawa and P. Ugliengo, *Chem. Mater.* 15, 3996–4004 (2003). doi:10.1021/cm0342804
- [103] C. Gatti, V.R. Saunders and C. Roetti, *J. Chem. Phys.* 101, 10686–10696 (1994). doi:10.1063/1.467882
- [104] S. Reine, E. Tellgren, A. Krapp, T. Kjærgaard, T. Helgaker, B. Jansik, S. Høst and P. Salek, *J. Chem. Phys.* 129, 104101 (2008). doi:10.1063/1.2956507
- [105] J.P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.* 77, 3865–3868 (1996). doi:10.1103/PhysRevLett.77.3865