PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2025, **27**, 23013

Combining the maximum overlap method with multiwavelets for core-ionisation energy calculations

Niklas Göllmann, [ac] Matthew R. Ludwig, [b] Peter Wind, [c] Laura E. Ratcliff *[bc] and Luca Frediani [c]

We present a protocol for computing core-ionisation energies for molecules, which is essential for reproducing X-ray photoelectron spectroscopy experiments. The electronic structures of both the ground state and the core-ionised states are computed using Multiwavelets and Density-Functional Theory, where the core-ionisation energies are computed by virtue of the ΔSCF method. To avoid the collapse of the core-hole state or its delocalisation, we make use of the Maximum Overlap Method, which provides a constraint on the orbital occupation, while avoiding the use of pseudopotentials. Combining Multiwavelets with the Maximum Overlap Method allows for the first time an all-electron calculation of core-ionisation energies with Multiwavelets, avoiding known issues connected to the use of Atomic Orbitals (slow convergence with respect to the basis set limit, numerical instabilities of core-hole states for large systems). We show that our results are consistent with previous Multiwavelet calculations which made use of pseudopotentials, and are generally more precise than corresponding Atomic Orbital calculations. We analyse the results in terms of precision compared to both Atomic Orbital calculations and Multiwavelets + pseudopotentials calculations. Moreover, we demonstrate how the protocol can be applied to target molecules of relatively large size. Both closed-shell and open-shell methods have been implemented.

Received 23rd April 2025,
Accepted 27th September 2025

DOI: 10.1039/d5cp01544h

rsc.li/pccp

1. Introduction

Core X-ray photoelectron spectroscopy (XPS) is a powerful technique for probing the electronic structure of materials, from molecules to surfaces to solids. By providing a direct measurement of core Binding Energies (BEs), that is the energy required to remove a particular core electron from the material, it is able to probe the local electronic structure, while offering insights into bonding nature and chemical and coordination environments. For valence XPS, there is a long-standing tradition of using theory, often based on density-functional theory (DFT)[1,2] calculations, to interpret experimental spectra. For core XPS, this has not traditionally been the case, with interpretation instead typically relying on comparisons to reference spectra. This is challenging, if not impossible, for core spectra of more complex materials, where there may be many overlapping peaks and no clear way to assign a given peak to a particular atom. This has motivated an increasing interest in methods for simulating core BEs, opening up the possibility of gaining new insights into complex materials using core XPS.

A range of approaches exist for calculating core BEs using DFT, including Koopmans' (which is rigorously defined for Hartree Fock only, but commonly employed for DFT), the $Z+1$ or equivalent core model[3,4] and ΔSCF.[5] Of these, ΔSCF is by far the most popular, and has been successfully applied to an increasingly wide range of systems.[6–8] Due to the challenges involved with calculating absolute BEs, it is common practice to align theoretical results with respect to experimental spectra. In such cases, relative core BEs are therefore sufficient for aiding in peak assignments and thus interpreting experimental results. However, ΔSCF has also shown promise for calculating absolute core BEs.[9,10] An alternative approach to DFT is GW,[11] which has in recent years also been applied to core BE calculations, *e.g.* as in ref. 12–14. Through a set of benchmark calculations to a test set of 65 molecules, covering C, N, O and F 1 *s* and also including relativistic effects, ΔSCF and GW were shown to have similarly high accuracies with respect to experiment.[12] Beyond DFT and GW, other quantum chemical approaches haves also been employed, examples of which

[a] *University of Münster, Organisch-Chemisches Institut and Center for Multiscale Theory and Computation, Corrensstraße 36, 48149 Münster, Germany*
[b] *Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, UK. E-mail: laura.ratcliff@bristol.ac.uk*
[c] *Hylleraas Centre, Department of Chemistry, UiT The Arctic University of Norway, N-9037 Tromsø, Norway*

include various coupled-cluster based approaches,[15–25] active space approaches,[26–28] configuration interaction,[29–31] Møller–Plesset perturbation theory[17,24,32–34] and algebraic–diagrammatic construction.[35–38] Recent works combining quantum embedding approaches with core–valence separation equation-of-motion coupled cluster approach have enabled core ionisation energy calculations of larger systems, including small molecules adsorbed on surfaces[39] and molecules with more than 100 atoms,[40] however, such quantum chemical approaches are typically limited to relatively small system sizes due to the high computational cost.

Like most molecular DFT calculations, the majority of ΔSCF calculations in gas phase have typically employed Gaussian basis sets. However, convergence with respect to basis set size is slow, with smaller basis sets typically not being flexible enough to treat core-excited states,[41] so that there is poor error cancellation between the ground-state and core-hole calculations. Accurate core BE calculations in Gaussian basis sets therefore usually require large basis sets, adding to the computational cost. Alternative strategies based on *e.g.* supplementing a basis set with additional functions in an approach akin to the $Z + 1$ approximation have been shown to provide a significant increase in accuracy without requiring too many additional basis functions.[42] However, a more general way to overcome the limitations of Gaussian basis sets and reach high precision is to use a systematic basis set. One such alternative that has gained a prominent role in the past few years is constituted by multiwavelets (MWs).[43–49] Multiwavelets are a specific realisation of wavelets, which use a set of polynomials on an interval.[50,51] The basis can be refined systematically and adaptively, providing strict error control on energetics and molecular properties. The comparatively large memory footprint of such a method has prevented its widespread use in the past, but modern high-performance computing (HPC) architectures combined with efficient implementation has lifted this limitation except for extremely large systems.[52] Moreover, several studies have shown that MWs become competitive in performance if high precision is requested.[52–54]

Another challenge of ΔSCF calculations is that of converging the core hole. Without constraints, the core hole may either delocalise or 'hop' between core orbitals of other atoms of the same species,[55,56] particularly when those orbitals have similar energies, *e.g.* due to having the same local chemical environment. This can lead to poor convergence, or indeed convergence to the wrong solution, and thus ultimately significant errors in core BEs. This combination of challenges motivated previous work in which the ΔSCF approach was implemented in a MW framework.[57] In that work, a combination of an all electron (AE) and pseudopotential (PSP) based approach was used, wherein the core states of the core-excited atom were treated explicitly, with the remainder of the atoms treated at the PSP level. The elimination of all other core orbitals prohibits the possibility of core-hole hopping, while the automatic refinement of the MW approach adapts to provide a more refined (accurate) grid where needed, thereby enabling accurate core BE calculations.

In this work, we combine MWs with an alternative approach to ensure the core hole remains sufficiently localised, namely the Maximum Overlap Method (MOM).[58] MOM, as well as its variant Initial Maximum Overlap Method (IMOM),[59] aim to keep the core hole localised on the correct atom by maximising the overlap between orbitals. By combining MOM and MWs for the first time, we demonstrate the possibility of achieving both high numerical precision and robust convergence for ΔSCF-based core BE calculations of molecules, while avoiding the use of pseudopotentials. We are aware that the ΔSCF method has limitations of its own: it is, for example, only capable of computing the main ionisation line, and not the so called shake-up satellite states, for which it is necessary to go beyond mean field methods like DFT.[60,61] We stress however that the main scope of this paper is to extend the availability of the ΔSCF method to AE MW calculations, rather than improving on the ΔSCF method itself.

We first provide an overview of the theory, including MWs, ΔSCF, and MOM and IMOM. We then present results for small molecules, amino acids and a large molecule, including comparisons with both Gaussian basis sets and the combined AE/PSP MW approach. Finally, we finish with a summary and conclusion.

2. Theory

2.1. Multiwavelets

Wavelets and MWs are a family of functions first constructed in the 1980s, and initially designed for signal processing.[62] They are localised both in real and Fourier space,[63] which makes them ideal for obtaining compact representations, which in turn lead to fast algorithms. It was about 20 years ago that their potential was first realised for electronic structure calculations,[64] leading to several practical implementations. For AE calculations, the first code to demonstrate their potential was MADNESS,[65] followed by MRChem[44,66] a few years later. To date, these are still the only two codes worldwide which are capable of AE MW electronic structure calculations for energies and molecular properties. They are both based on Alpert's realization of MWs, which in practice define the mother scaling functions as a set of polynomials $\varphi_j, j = 0,k$ up to order k on an interval.[50] Such polynomial functions can then be dilated and translated to obtain progressively finer representations on a grid:

$$\varphi_{jl}^{n} = 2^{n/2}\varphi_j(2^n x - l)$$

where n is the current scale and l is the translation index ($l = 0 \ldots 2^n - 1$). This defines a family of nested vector spaces $\ldots V_{n-1} \subset V_n \subset V_{n+1} \ldots$ which is dense in L^2. Wavelet functions ψ_{jl}^{n} are obtained as the orthogonal complement of two successive scaling spaces $W_n = V_{n+1} \ominus V_n$. Multidimensional representations are then obtained as standard tensor products, to represent *e.g.* orbitals and the electronic density.

For a practical realisation of an AE code, several additional advances were necessary: an adaptive representation of

functions,[51] which allows grids to be refined locally, where needed; the non-standard form of operators,[67] which is sparse and enables adaptivity at the stage of operator application; a separated representation of operators, which mitigates the "curse of dimensionality" by reducing the prefactor of operator applications from k^{2d} to k^{d+1}, where d is the dimensionality of the system[68]; the integral formulation of the self-consistent field (SCF) equations[69,70] coupled with a Krylov-accelerated inexact Newton (KAIN) method for faster convergence;[71] and specific organisation of the huge amount of data, that describes a set of orbitals in the computational code, to allow efficient parallel processing on HPC systems.[52]

2.2. Core binding energy calculations

The simplest approach to the calculation of core BEs within DFT is the Koopmans' approach, which uses the negative of the core KS orbital energies as an approximation to the core BEs. While this has been shown to give good results for relative core BEs for some systems, it does not, however, take into account final state effects. For systems where final state effects are important, a better approach is ΔSCF, in which an explicit core hole is introduced to a given atom and orbital of interest. The associated BE is then calculated using $\mathrm{BE} = E^{N-1}_{\mathrm{final}} - E^{N}_{\mathrm{initial}}$, where E^{N-1}_{final} is the energy of the system in its final state, *i.e.* with an explicit core hole, and E^{N}_{initial} is the initial ground-state energy.

Computing core-ionisation energies brings about the challenge of stabilising the core-hole state: a standard SCF optimisation would not work, because the hole state would likely migrate towards the valence region, resulting in a standard calculation of the ionic system. Several strategies to lock the core-hole state have been devised:

- Broken-symmetry guess: a starting guess for the core-hole calculation is obtained by performing a ground-state calculation with a small additional charge at the core site, as in *e.g.* ref. 10. The resulting orbitals are then used as a guess for the core-hole calculation. While this symmetry breaking can help reduce core-hole hopping, no further constraints are applied, so it does not necessarily prevent the collapse of the core state.
- $Z + 1$ technique: instead of the core-hole calculation, one performs a calculation of a system with an augmented nuclear charge at the core site. This method has better stability, but the results are less accurate than ΔSCF because the system is significantly perturbed. Alternatively, $Z + 1$ could be used to generate a better input guess, but like with the broken symmetry guess, if no further constraints are applied, there may still be convergence problems.
- Pseudopotentials: all core electrons apart from those associated with the target atom are described *via* PSPs. The core-hole state is thus well separated in energy from all other states, preventing core-hole hopping and stabilising convergence.
- MOM and IMOM: the optimisation is constrained by keeping the maximum possible overlap between both the core-hole and other orbitals with their counterparts from either the previous iteration along the optimisation (MOM), or the starting guess of the calculation (IMOM).

The first two methods are the simplest to adopt as they require minimal effort to modify the implementation of ΔSCF. The last two are more general. The use of PSPs in combination with MWs has been demonstrated in a previous study.[72] In the present work we combine MWs with both MOM and IMOM.

2.3. MOM and IMOM

The goal of a ΔSCF calculation is to optimise a non-Aufbau occupation using the SCF procedure. In most cases, however, the SCF procedure is employed to find the solution for an Aufbau occupation. The challenge posed by core-hole calculations is constituted by the instability of the core-ionised state which could easily collapse, yielding a valence-ionisation state, which has a much lower energy. Additionally, the core-hole could switch site or even delocalise, rendering the results useless when compared to experimental measurements.[55,56]

To stabilise the core-hole state, MOM[58] attempts to maintain the largest possible overlap between the orbitals at a given iteration in the SCF optimisation and the previous iterate. This is achieved by computing the overlap matrix $\mathbf{O}$ between the two sets of orbitals at successive iterations k:

$$O^{(k-1,k)}_{ir} = \langle \psi^{(k-1)}_i | \psi^{(k)}_r \rangle, \tag{1}$$

where the indices i and r denote doubly occupied orbitals and all orbitals, respectively. For each orbital, a weight measure is computed by taking the norm of the corresponding column in the overlap matrix:

$$p_r = \left(\sum_i^n O^2_{ir} \right)^{1/2} \tag{2}$$

The p_r values are then used to assign the occupation in the next iteration, in such a way that the core-hole orbitals stay as close as possible to the previous iterate. This helps prevent both variational collapse and hopping or delocalisation. The IMOM[59] procedure is essentially identical. The only difference is that the reference orbitals to compute the weights p_r are always the initial guess orbitals, rather than being updated at every iteration,

$$O^{(0,k)}_{ir} = \langle \psi^{0}_i | \psi^{k}_r \rangle. \tag{3}$$

This, in theory, should decrease the likelihood that a variational collapse will occur.[59]

2.4. Computational details

For both MADNESS and MRChem calculations, ground-state calculations employed localised orbitals, while core-hole calculations employed canonical orbitals. The ground-state orbitals were used as an initial guess for the core-hole calculations for both codes. Calculations were performed using the PBE functional.[73] Unless otherwise stated, MRChem calculations employed a world precision of 10^{-5} (abbreviated as MW5), and all results are obtained using the IMOM protocol, as preliminary calculations showed rare issues with respect to variational collapse using the MOM method. MADNESS calculations employed the mixed AE/PSP protocol, using HGH-GTH PSPs.[74,75]

The MADNESS ground-state calculations used a precision threshold of 10^{-4} followed by 10^{-6} (polynomial orders $k = 6$ and $k = 8$ respectively), while core-hole calculations directly used a precision threshold of 10^{-6} (wavelet order $k = 8$). Both the density and wave function residuals employed convergence criteria of 10^{-3}. Gaussian basis set calculations were performed using the NWChem code.[76] For each core-hole calculation, an input guess was constructed by performing a ground-state calculation in which a fictitious charge of 0.1 e was added to the target atom, following ref. 10. Calculations were performed using the def2 basis set family.[77] Spin restricted calculations were performed for all systems, while for glycine and alanine unrestricted calculations were also performed. The molecular structures of ethanol and vinyl fluoride were relaxed using MRChem in the restricted formalism, using localised orbitals with a world precision of 10^{-6}. Spin unrestricted calculations were also performed for vinyl fluoride geometries with a range of C–C bond lengths, constructed from the relaxed structure by translating half the molecule to achieve the target separation. Structures for glycine and alanine were extracted from their respective crystal structures and geometry optimised using BigDFT[78] with a grid spacing of 0.5 Bohr and coarse (fine) radius multipliers of 5 (7), in all cases using PBE. The atomic structure for 2CzPN was the relaxed structure taken directly from ref. 79.

For the purpose of the present work, we have not explicitly analysed the computational performance nor the memory footprint. However, ΔSCF calculations are conceptually very simple requiring only single-point energy calculations. We have previously shown that MW methods can be made to scale linearly with the system size for pure density functional theory (DFT) methods, slightly worse than linearly ($N^{1+\alpha}$ with $\alpha \simeq 0.3$) if exact exchange is included.[80] On the other hand, the protocol to compute core-ionised states requires the use of canonical orbitals, yielding a quadratic scaling. Linear scaling could still be achieved by computing relevant quantities in the localised basis before transforming back to the canonical one. However, this procedure was not implemented for the present study.

3. Results

In the following, we present results for a range of system sizes, from small molecules, where we demonstrate the ability of the MW framework to achieve very high precision, through to a large molecule, where we show both the ability to treat large systems and the robustness of the combined MW IMOM approach, even for systems with a large number of atoms of the same species.

3.1. Small molecules

For our first tests, we take ethanol and vinyl fluoride, chosen for being very small molecules, which nonetheless have two distinct C environments. Table 1 shows the ground-state and core-hole energies for the two molecules, calculated using the MW IMOM approach for a series of world precisions, as well as using the def2 basis set family. Fig. 1 shows the corresponding BEs, and also includes results for the MW AE/PSP approach. Compared to the most precise MW IMOM calculation, the def2 total energies differ significantly, by around 5–7 eV for the double zeta basis, decreasing to at most 0.14 eV for the quadruple zeta basis. However, as expected, the BEs benefit from error cancellation, with the difference with respect to the MW IMOM results decreasing to around 1.5 eV for the double zeta basis and less than 0.08 eV for the quadruple zeta basis. These errors are very similar across C environments in both molecules, so that the error in relative BE is less than 0.1 eV even for the double zeta basis compared to the most precise MW IMOM results (see Table S1 and Fig. S1 in the SI). The two different MW approaches show excellent agreement, with the AE/PSP BEs results differing from the highest precision IMOM

Table 1 Comparison between total energies for ground state and core hole calculations for ethanol and vinyl fluoride, calculated for the def2 basis set family and a series of world precisions for the MW IMOM approach. MW AE/PSP energies are omitted since the use of PSPs precludes a direct comparison. Differences are reported relative to MW IMOM results using a world precision of 10^{-8}. All values are in eV

	Ground state	Δ_{GS}	C_1	Δ_{C_1}	C_2	Δ_{C_2}
Ethanol						
def2-SVP	−4210.34083260	5.31245782	−3910.69896692	6.68120313	−3912.12437184	6.74606783
def2-TZVP	−4215.24887563	0.40441479	−3916.64107365	0.73909640	−3918.12982277	0.74061690
def2-QZVP	−4215.59047745	0.06281297	−3917.23880636	0.14136369	−3918.75773216	0.11270751
Multiwavelet (initial maximum overlap method) 10^{-4}	−4215.64961530	0.00367513	−3917.37695587	0.00321417	−3918.86487573	0.00556394
MW (IMOM) 10^{-5}	−4215.65322617	0.00006426	−3917.38053712	−0.00036708	−3918.87033898	0.00010069
MW (IMOM) 10^{-6}	−4215.65328860	0.00000183	−3917.38086218	−0.00069214	−3918.87038454	0.00005513
MW (IMOM) 10^{-7}	−4215.65329010	0.00000032	−3917.38038926	−0.00021922	−3918.87041800	0.00002167
MW (IMOM) 10^{-8}	−4215.65329043	—	−3917.38017004	—	−3918.87043967	—
Vinyl fluoride						
def2-SVP	−4829.83770206	6.01642799	−4529.01676646	7.49937180	−4531.33997522	7.52241717
def2-TZVP	−4835.50767314	0.34645691	−4535.83728200	0.67885626	−4538.17714577	0.68524662
def2-QZVP	−4835.79935913	0.05477092	−4536.38514167	0.13099659	−4538.75373277	0.10865962
MW (IMOM) 10^{-4}	−4835.84606267	0.00806738	−4536.51086111	0.00527714	−4538.85693569	0.00545670
MW (IMOM) 10^{-5}	−4835.85394889	0.00018116	−4536.51609614	0.00004212	−4538.86222573	0.00016666
MW (IMOM) 10^{-6}	−4835.85412681	0.00000325	−4536.51584129	0.00029697	−4538.86234967	0.00004272
MW (IMOM) 10^{-7}	−4835.85412897	0.00000108	−4536.51612946	0.00000880	−4538.86238553	0.00000686
MW (IMOM) 10^{-8}	−4835.85413005	—	−4536.51613826	—	−4538.86239239	—

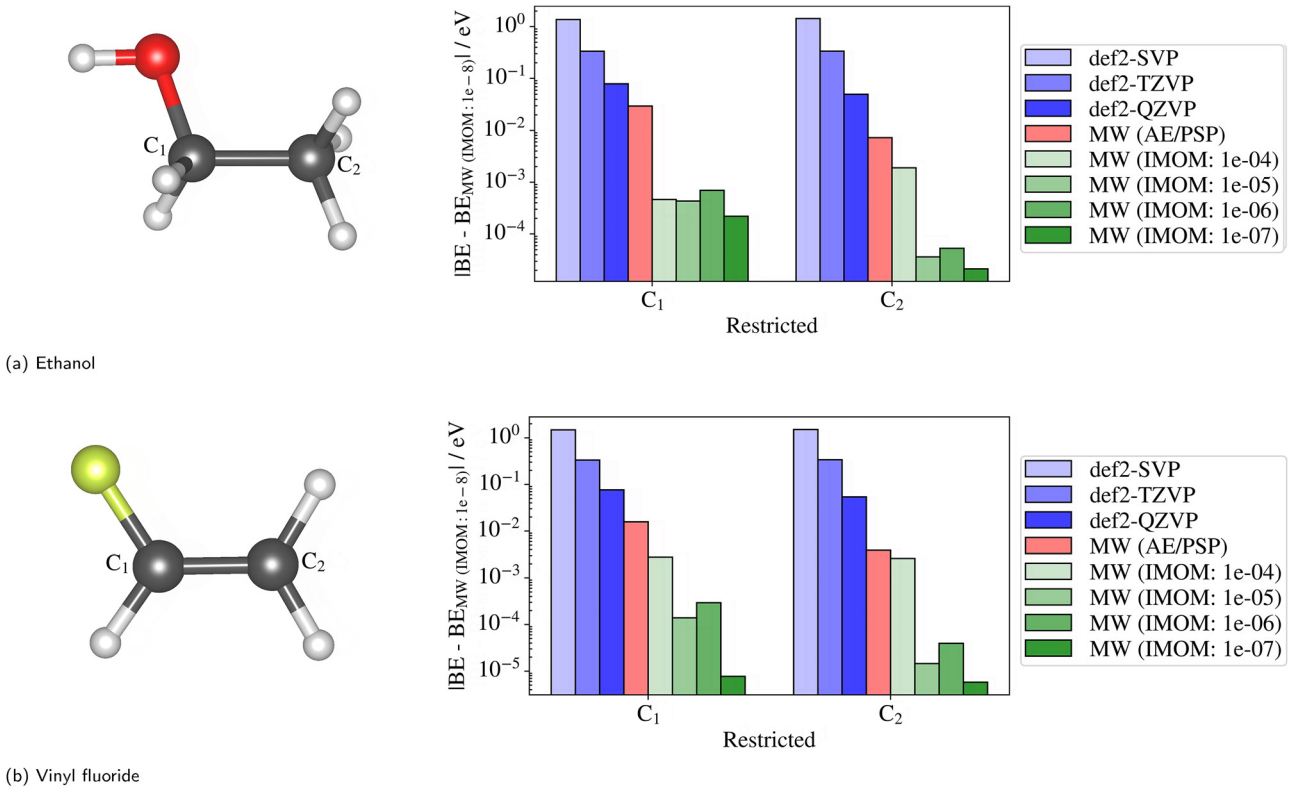

Fig. 1 Comparison of calculated absolute BEs for ethanol and vinyl fluoride, for the def2 family of basis sets, MW calculations using the AE/PSP implementation in MADNESS, and the MW IMOM implementation in MRChem, where the latter has been calculated for a series of different world precisions. Results are given relative to the MW IMOM implementation in MRChem with a world precision of 10^{-8}. Shown on the left are the employed atomic structures, labelled with the different atomic environments. See Table S1 in the SI for tabulated values.

results by less than 0.03 eV, around the level of thermal broadening effects at room temperature, and well below typical XPS resolution. In fact, for absolute BEs, both MW approaches as well as def2-QZVP values all agree to within 0.08 eV, which is less than half the experimental resolution which might be expected for synchrotron hard X-ray photoelectron spectroscopy (HAXPES) experiments, while the relative BEs all agree to within 0.03 eV.

Considering now the convergence with respect to world precision for the MW IMOM results, the BEs do not show the expected systematic convergence behaviour. Looking at the total energies, the ground-state energies converge smoothly, while the core-hole energies are less systematic. One possibility is that core-hole calculations are not strictly bound from below like the ground state and this might cause instabilities. Although IMOM prevents the full collapse of the core-hole state it might still not be able to avoid such small artifacts. This is *e.g.* particularly evident for the C1-hole state of ethanol, where lower precision results (MW5, MW6 and MW7) display lower total energies than MW8 precision. Some deviations might be expected where *e.g.* a lower precision calculation might by chance more than satisfy the precision requirements, however it is not clear why the core-hole calculations are so susceptible to showing non-systematic convergence, while the ground-state calculations are not. Additional calculations were performed using canonical ground-state orbitals, and similar behaviour was observed.

This behaviour affects the error on the BE, which is now a difference between two absolute energy errors. As a result, the lack of systematicity is amplified, leading to a somehow erratic convergence to the MW8 result. Nonetheless, these differences are very small, with the total energies differing by around 0.01 eV even for the least precise calculations (10^{-4} world precision), decreasing to the order of 10^{-4} eV for the next smallest precision (10^{-5}). These differences are smaller than any of the differences with respect to the other approaches, and more than sufficient for practical calculations. Therefore, a world precision of 10^{-5} is used for all subsequent calculations.

Finally, we also use vinyl fluoride as an example of how the different approaches compare for a system out of equilibrium. To this end, Fig. 2 shows the relative BEs of vinyl fluoride across a range of C–C bond lengths, calculated in the unrestricted formalism. All approaches show qualitatively similar behaviour up to 2 Å, while the relative BEs agree to within around 0.02 eV for all but the def2-SVP basis set. For bond lengths beyond 2 Å, all approaches start to break down, either failing to converge within a reasonable number of steps or showing large jumps in energy despite only a small change in bond lengths. However, for structures far from equilibrium, single determinant methods are not expected to give accurate results, irrespective of the basis

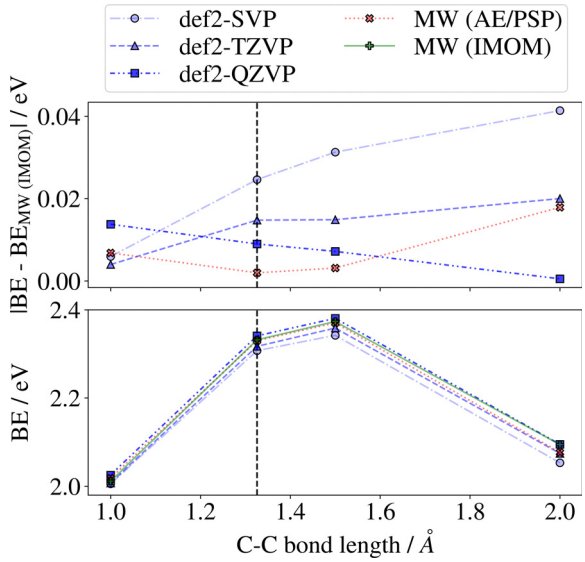

Fig. 2 Comparison of calculated relative BEs for vinyl fluoride with varying C–C bond length, for the def2 family of basis sets, MW calculations using the AE/PSP implementation in MADNESS, and the MW IMOM implementation in MRChem in the unrestricted formalism. The top panel also shows results relative to the MW IMOM implementation in MRChem. See Table S2 in the SI for tabulated values.

set or other implementation details, and therefore an alternative approach should be used.

3.2. Amino acids

As our second example, we take the amino acids glycine (Gly) and alanine (Ala), which have been well characterised from both a theoretical and experimental perspective, including previous calculations using the AE/PSP approach implemented in ref. 57. In this work, we take a single conformer, and perform calculations in both restricted and unrestricted formalisms, again comparing MW IMOM, MW AE/PSP and def2 basis set calculations. Results are depicted in Fig. 3, with the absolute BE given in Table S3 in the SI.

Similar to the small molecules, the double zeta BEs differ significantly from the MW IMOM values, by 1–2 eV. The error is already much smaller for the triple zeta basis, with an average difference of 0.4 eV across all calculations, further reducing to 0.07 eV for the quadruple zeta basis. The difference between MW approaches is again small, with an average difference of 0.04 eV between the two. The relative BEs again benefit from error cancellation. There are some variations between different core states, but in the worst case, the relative BEs of the triple zeta differ by 0.06 eV with respect to the MW IMOM results, while the quadruple zeta and AE/PSP results differ by even less (see Fig. S2 in the SI).

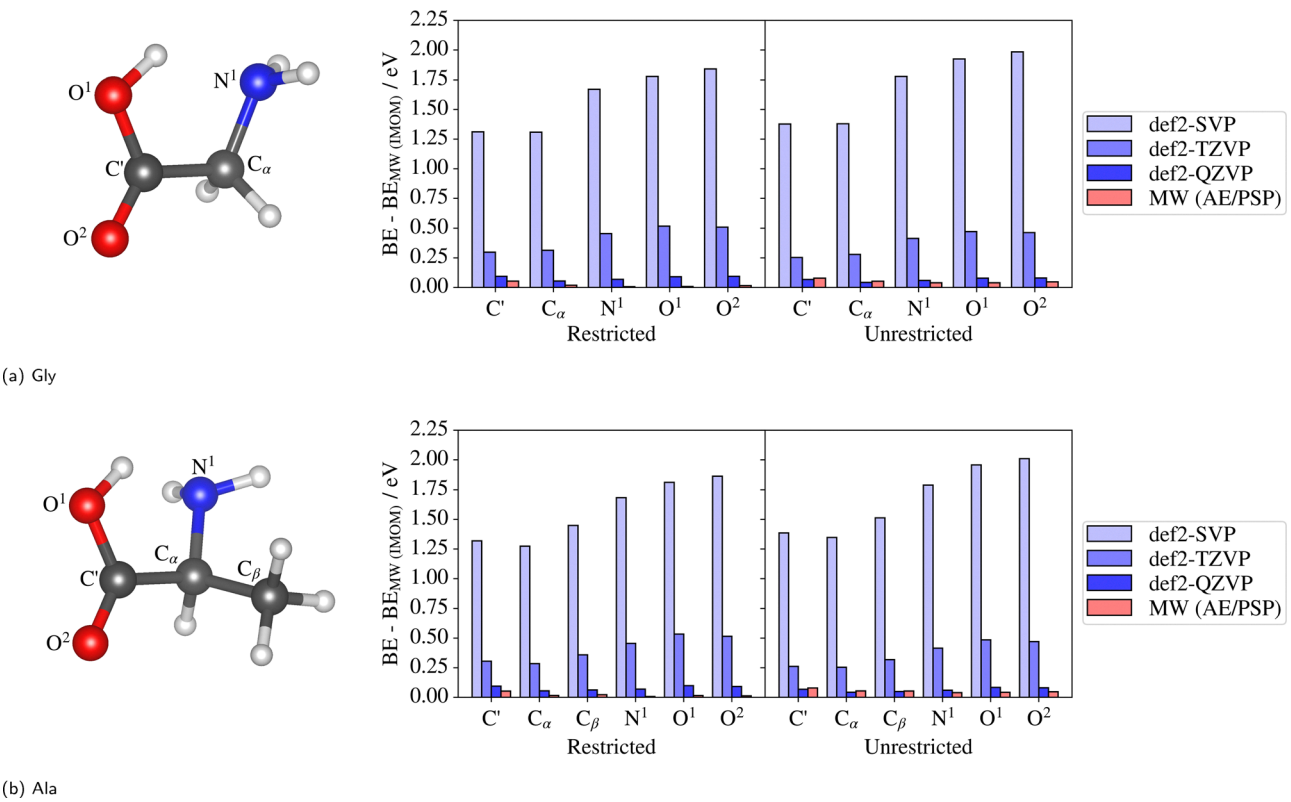

Fig. 3 Comparison of calculated absolute BEs for the amino acids glycine and alanine for both restricted and unrestricted calculations, for the def2 family of basis sets, and the two different MW implementations. Results are given relative to the MW IMOM implementation in MRChem. Shown on the left are the atomic structures of the employed conformers, labelled with the different atomic environments.

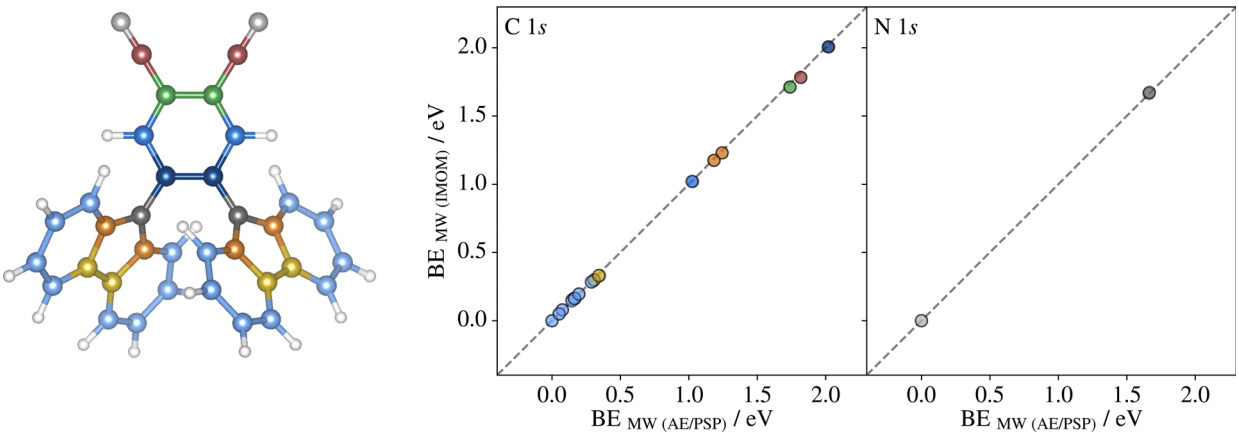

Fig. 4 Comparison between relative BEs for the relaxed 2CzPN molecule, for the MW IMOM implementation in MRChem and the mixed AE/PSP MW MADNESS results from ref. 79. Shown on the left is the atomic structure of 2CzPN, where the atoms have been coloured according to their different local chemical environments, where the two N atoms are in shades of grey, and the C atoms are in a range of colours. The corresponding values are given in Table S4 in the SI.

Finally, using a restricted approach results in a large shift in the absolute BEs, of around 7–9 eV depending on the chemical species. However, the effect on the relative BEs is much smaller, with the mean absolute deviation between restricted and unrestricted relative BEs being less than 0.01 eV for all basis sets. Thus, the error in relative BE introduced by using a spin restricted formalism is lower than the basis set effects.

3.3. 2CzPN

We finish with the example of 2CzPN, a prototypical thermally activated delayed fluorescence (TADF)-based organic light emitting diode (OLED) emitter. 2CzPN has previously been investigated with MADNESS,[79] where it was shown that theory is essential for interpreting the experimental spectra, with three distinct peaks in the C 1s spectra coming from seven underlying chemical environments. Being a larger molecule (54 atoms/119 doubly-occupied orbitals), it is also a good test of both the ability of MRChem to treat large systems, and of the effectiveness of the IMOM approach in the case when there are a large number of C atoms, in some cases with the same local chemical environment, and thus a high likelihood of core-hole hopping occurring.

All MW IMOM calculations successfully converged, demonstrating the ability of IMOM to handle C atoms with multiple similar local chemical environments. Shown in Fig. 4 are the core BEs for the relaxed 2CzPN structure, where the MW IMOM results are compared with the mixed AE/PSP MW results from previous work.[79] As can be seen, the relative BEs for each approach are in excellent agreement, for both C and N atoms. Furthermore, the absolute BEs have a mean absolute deviation between the two approaches of 0.02 eV. As discussed above, this is of the order of thermal broadening effects at room temperature and well below experimental resolution. In short, the MW IMOM MRChem approach proves to be robust and accurate for core BEs calculations of large molecules.

4. Conclusions

In this work, we have combined the Maximum Overlap Method and Initial Maximum Overlap Method with a Multiwavelet approach to describe the electronic structure of a molecule. MOM and IMOM are designed to stabilise the convergence of core-hole state calculations, without the use of pseudopotentials, by ensuring the hole state is kept as close as possible to the original electronic state. Multiwavelets, on the other hand, provide a robust and systematic approach to the complete basis set (CBS) limit, overcoming limitations of atomic orbital (AO) bases. This is especially critical for core properties, which negate one of the basic assumptions behind using AOs: error cancellation, which generally mitigates an inadequate description of core states, but it can be exploited when electronic structure changes are essentially affecting the valence electrons. This combination has proven very successful: we are able to reproduce results obtained using an approach which combines MWs with the use of pseudopotentials for all but one core-excited atom (for which all electrons are treated explicitly) and to a precision which is at least one order of magnitude better than the resolution of XPS experiments. We show that such a precision is achieved consistently, albeit showing a less systematic trend with respect to increasing precision than expected. We are also able to treat a large molecule with many chemically similar atoms, which is a notorious challenge for this kind of calculation. Between MOM and IMOM, the latter has proved slightly superior in terms of reliability: although both are implemented, all results featured in the paper are obtained with IMOM, as the MOM procedure can show a variational collapse in rare cases. We have designed a robust protocol which combines the use of localised orbitals in the ground state, which allows one to select a core orbital located at a specific atom, with canonical orbitals for the excited state, which enables convergence to the requested core-hole state by decoupling SCF equations during the optimisation. Together with IMOM, this effectively prevented the collapse of the core hole for all cases investigated. Work is

ongoing to extend the current work to include Hartree-Fock (HF) exchange. Thereafter, we would like to apply the method to a range of larger molecules, including, for example, extending previous work on exploiting the local nature of core XPS for probing the effects of disorder in 2CzPN to other TADF emitters. The complexity of the measured spectra for such molecules necessitates the use of theory to enable the interpretation of experimental results, while the large size of such molecules requires a robust and efficient computational approach, which our current approach provides.

Conflicts of interest

There are no conflicts of interest to declare.

Data availability

The data supporting this article have been included in the supplementary information (SI). Supplementary information contains additional plots of relative BEs for the small molecules and amino acids, tabulated data underlying the figures presented in the main paper, and a schematic showing the employed atom labelling for 2CzPN. See DOI: **https://doi.org/10.1039/d5cp01544h**.

Acknowledgements

We are grateful for computational support from the UK national high performance computing service, ARCHER2, for which access was obtained *via* the UKCP consortium and funded by EPSRC grant ref EP/X035891/1. MRL acknowledges support from the Engineering and Physical Sciences Research Council (EP/W524414/1). LER thanks Anna Regoutz for useful discussions. The financial support from the Research Council of Norway through its Centres of Excellence scheme (Hylleraas centre, 262695) and the FRIPRO scheme (ReMRChem, 324590) is acknowledged. The support from Sigma2, The National Infrastructure for High Performance Computing and Data Storage in Norway, under the NN14654K grant of computer time is also acknowledged.

Notes and references

1 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.
2 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
3 W. L. Jolly and D. N. Hendrickson, *J. Am. Chem. Soc.*, 1970, **92**, 1863–1871.
4 P. S. Bagus, C. Sousa and F. Illas, *Phys. Chem. Chem. Phys.*, 2020, **22**, 22617–22626.
5 P. S. Bagus, *Phys. Rev.*, 1965, **139**, A619–A634.
6 F. Viñes, C. Sousa and F. Illas, *Phys. Chem. Chem. Phys.*, 2018, **20**, 8403–8410.
7 N. A. Besley, A. T. B. Gilbert and P. M. W. Gill, *J. Chem. Phys.*, 2009, **130**, 124308.
8 D. Hait and M. Head-Gordon, *J. Phys. Chem. Lett.*, 2021, **12**, 4517–4529.
9 N. Pueyo Bellafont, F. Viñes and F. Illas, *J. Chem. Theory Comput.*, 2016, **12**, 324–331.
10 J. M. Kahk and J. Lischner, *Phy. Rev. Mater.*, 2019, **3**, 100801.
11 L. Hedin, *Phys. Rev.*, 1965, **139**, A796–A823.
12 D. Golze, L. Keller and P. Rinke, *J. Phys. Chem. Lett.*, 2020, **11**, 1840–1847.
13 D. Mejia-Rodriguez, A. Kunitsa, E. Aprà and N. Govind, *J. Chem. Theory Comput.*, 2021, **17**, 7504–7517.
14 J. Li, Y. Jin, P. Rinke, W. Yang and D. Golze, *J. Chem. Theory Comput.*, 2022, **18**, 7570–7585.
15 S. Sen, S. Avijit and D. Mukherjee, *Mol. Phys.*, 2013, **111**, 2625–2639.
16 S. Coriani and H. Koch, *J. Chem. Phys.*, 2015, **143**, 181103.
17 A. Holme, K. J. Børve, L. J. Sæthre and T. D. Thomas, *J. Chem. Theory Comput.*, 2011, **7**, 4104–4114.
18 J. Liu, D. Matthews, S. Coriani and L. Cheng, *J. Chem. Theory Comput.*, 2019, **15**, 1642–1651.
19 M. L. Vidal, A. I. Krylov and S. Coriani, *Phys. Chem. Chem. Phys.*, 2020, **22**, 2693–2703.
20 X. Zheng and L. Cheng, *J. Chem. Theory Comput.*, 2019, **15**, 4945–4955.
21 F. Uhl and V. Staemmler, *J. Electron Spectrosc. Relat. Phenom.*, 2019, **233**, 90–96.
22 D. A. Matthews, *Mol. Phys.*, 2020, **118**, e1771448.
23 J. E. Arias-Martinez, L. A. Cunha, K. J. Oosterbaan, J. Lee and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2022, **24**, 20728–20741.
24 A. Morgunov, H. K. Tran, O. R. Meitei, Y.-C. Chien and T. Van Voorhis, *J. Phys. Chem. A*, 2024, **128**, 6989–6998.
25 B. Peng and K. Kowalski, *J. Chem. Theory Comput.*, 2018, **14**, 4335–4352.
26 P. S. Bagus, C. Sousa and F. Illas, *Theor. Chem. Acc.*, 2019, **138**, 61.
27 C. Sousa, P. S. Bagus and F. Illas, *J. Phys. Chem. A*, 2024, **128**, 895–901.
28 M. Huang and F. A. Evangelista, *J. Chem. Theory Comput.*, 2024, **20**, 7990–8000.
29 S. Carniato, G. Dufour, Y. Luo and H. Ågren, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2002, **66**, 045105.
30 N. Galamba and B. J. C. Cabral, *J. Chem. Phys.*, 2018, **148**, 044510.
31 M. W. D. Hanson-Heine, M. W. George and N. A. Besley, *J. Chem. Phys.*, 2019, **151**, 034104.
32 J. Shim, M. Klobukowski, M. Barysz and J. Leszczynski, *Phys. Chem. Chem. Phys.*, 2011, **13**, 5703–5711.
33 J. V. Jorstad, T. Xie and C. M. Morales, *Int. J. Quantum Chem.*, 2022, **122**, e26881.
34 N. A. Besley, *J. Chem. Theory Comput.*, 2021, **17**, 3644–3651.
35 O. Plekan, V. Feyer, R. Richter, M. Coreno, M. de Simone, K. C. Prince, A. B. Trofimov, E. V. Gromov, I. L. Zaytseva and J. Schirmer, *Chem. Phys.*, 2008, **347**, 360–375.
36 C. E. V. d Moura and A. Y. Sokolov, *Phys. Chem. Chem. Phys.*, 2022, **24**, 4769–4784.
37 C. E. V. de Moura and A. Y. Sokolov, *J. Phys. Chem. A*, 2024, **128**, 5816–5831.

38 N. P. Gaba, C. E. V. d Moura, R. Majumder and A. Y. Sokolov, *Phys. Chem. Chem. Phys.*, 2024, **26**, 15927–15938.
39 R. A. Opoku, C. Toubin and A. S. Pereira Gomes, *Phys. Chem. Chem. Phys.*, 2022, **24**, 14390–14407.
40 B. Jangid, M. R. Hermes and L. Gagliardi, *J. Phys. Chem. Lett.*, 2024, **15**, 5954–5963.
41 A. E. A. Fouda and N. A. Besley, *Theor. Chem. Acc.*, 2017, **137**, 6.
42 M. W. Hanson-Heine, M. W. George and N. A. Besley, *Chem. Phys. Lett.*, 2018, **699**, 279–285.
43 R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan and G. Beylkin, *J. Chem. Phys.*, 2004, **121**, 11587–11598.
44 S. R. Jensen, S. Saha, J. A. Flores-Livas, W. Huhn, V. Blum, S. Goedecker and L. Frediani, *J. Phys. Chem. Lett.*, 2017, **8**, 1449–1457.
45 T. Yanai, G. I. Fann, Z. Gan, R. J. Harrison and G. Beylkin, *J. Chem. Phys.*, 2004, **121**, 2866–2876.
46 S. R. Jensen, T. Flå, D. Jonsson, R. S. Monstad, K. Ruud and L. Frediani, *Phys. Chem. Chem. Phys.*, 2016, **18**, 21145–21161.
47 A. Brakestad, S. R. Jensen, P. Wind, M. D'Alessandro, L. Genovese, K. H. Hopmann and L. Frediani, *J. Chem. Theory Comput.*, 2020, **16**, 4874–4882.
48 F. A. Bischoff, *Adv. Quantum Chem.*, 2019, **79**, 3–52.
49 N. Vence, R. Harrison and P. Krstic, *Phys. Rev. A: At., Mol., Opt. Phys.*, 2012, **85**, 033403.
50 B. K. Alpert, *SIAM J. Math. Anal.*, 1993, **24**, 246–262.
51 B. Alpert, G. Beylkin, D. Gines and L. Vozovoi, *J. Comput. Phys.*, 2002, **182**, 149–190.
52 P. Wind, M. Bjørgve, A. Brakestad, G. A. S. Gerez, S. R. Jensen, R. D. R. Eikås and L. Frediani, *J. Chem. Theory Comput.*, 2023, **19**, 137–146.
53 Q. Pitteloud, P. Wind, S. R. Jensen, L. Frediani and F. Jensen, *J. Chem. Theory Comput.*, 2023, **19**, 5863–5871.
54 M. Gubler, J. A. Finkler, S. R. Jensen, S. Goedecker and L. Frediani, *J. Phys. Chem. A*, 2025, **129**, 1469–1477.
55 P. S. Bagus and H. F. Schaefer, *J. Chem. Phys.*, 1972, **56**, 224–226.
56 I. Tolbatov and D. M. Chipman, *Theor. Chem. Acc.*, 2017, **136**, 82.
57 J. M. Pi, M. Stella, N. K. Fernando, A. Y. Lam, A. Regoutz and L. E. Ratcliff, *J. Phys. Chem. Lett.*, 2020, **11**, 2256–2262.
58 A. T. B. Gilbert, N. A. Besley and P. M. W. Gill, *J. Phys. Chem. A*, 2008, **112**, 13164–13171.
59 G. M. J. Barca, A. T. B. Gilbert and P. M. W. Gill, *J. Chem. Theory Comput.*, 2018, **14**, 1501–1509.
60 L. Cederbaum, *Mol. Phys.*, 1974, **28**, 479–493.
61 R. L. Martin and D. A. Shirley, *J. Chem. Phys.*, 1976, **64**, 3685–3689.
62 F. Keinert, *Wavelets and Multiwavelets*, 2003.
63 G. Strang, *Bull. Am. Math. Soc.*, 1993, **28**, 288–305.
64 K. Cho, T. A. Arias, J. D. Joannopoulos and P. K. Lam, *Phys. Rev. Lett.*, 1993, **71**, 1808–1811.
65 R. J. Harrison, G. Beylkin, F. A. Bischoff, J. A. Calvin, G. I. Fann, J. Fosso-Tande, D. Galindo, J. R. Hammond, R. Hartman-Baker, J. C. Hill, J. Jia, J. S. Kottmann, M.-J. Y. Ou, J. Pei, L. E. Ratcliff, M. G. Reuter, A. C. Richie-Halford, N. A. Romero, H. Sekino, W. A. Shelton, B. E. Sundahl, W. S. Thornton, E. F. Valeev, A. Vázquez-Mayagoitia, N. Vence, T. Yanai and Y. Yokoi, *SIAM J. Sci. Comput.*, 2016, **38**, S123–S142.
66 R. Bast, M. Bjorgve, R. Di Remigio, A. Durdek, L. Frediani, G. Gerez, S. R. Jensen, J. Juselius, R. Monstad and P. Wind, *MRChem: MultiResolution Chemistry*, **https://zenodo.org/doi/10.5281/zenodo.10522608**.
67 G. Beylkin, V. Cheruvu and F. Pérez, *Appl. Comput. Harmon. Anal.*, 2008, **24**, 354–377.
68 G. I. Fann, R. J. Harrison and G. Beylkin, *J. Phys.: Conf. Ser.*, 2005, **16**, 461.
69 M. H. Kalos, *Phys. Rev.*, 1962, **128**, 1791–1795.
70 S. R. Jensen, A. Durdek, M. Bjørgve, P. Wind, T. Flå and L. Frediani, *J. Math. Chem.*, 2023, **61**, 343–361.
71 R. J. Harrison, *J. Comput. Chem.*, 2004, **25**, 328–334.
72 L. E. Ratcliff, W. S. Thornton, Á. Vázquez-Mayagoitia and N. A. Romero, *J. Phys. Chem. A*, 2019, **123**, 4465–4474.
73 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
74 S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 1703–1710.
75 C. Hartwigsen, S. Goedecker and J. Hutter, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 3641–3662.
76 M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus and W. de Jong, *Comput. Phys. Commun.*, 2010, **181**, 1477–1489.
77 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297.
78 L. E. Ratcliff, W. Dawson, G. Fisicaro, D. Caliste, S. Mohr, A. Degomme, B. Videau, V. Cristiglio, M. Stella, M. D'Alessandro, S. Goedecker, T. Nakajima, T. Deutsch and L. Genovese, *J. Chem. Phys.*, 2020, **152**, 194110.
79 N. K. Fernando, M. Stella, W. Dawson, T. Nakajima, L. Genovese, A. Regoutz and L. E. Ratcliff, *Phys. Chem. Chem. Phys.*, 2022, **24**, 23329–23339.
80 P. Wind, M. Bjørgve, A. Brakestad, G. A. Gerez, S. R. Jensen, R. D. R. Eikås and L. Frediani, *J. Chem. Theory Comput.*, 2023, **19**, 137–146.