PCCP

ROYAL SOCIETY OF CHEMISTRY

PERSPECTIVE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Phys. Chem. Chem. Phys.*, 2022, **24**, 12787

Advances in the OCEAN-3 spectroscopy package

John Vinson

Received 2nd March 2022, Accepted 7th May 2022

DOI: 10.1039/d2cp01030e

rsc.li/pccp

The OCEAN code for calculating valence- and core-level spectra using the Bethe-Salpeter equation is briefly reviewed. OCEAN is capable of calculating optical absorption, near-edge X-ray absorption or non-resonant scattering, and resonant inelastic X-ray scattering, requiring only the structure of the material as input. Improved default behavior and reduced input requirements are detailed as well as new capabilities, such as incorporation of final-state-dependent broadening, finite-temperature dependence, and flexibility in the density-functional theory exchange–correlation potentials. OCEAN is built on top of a plane-wave, pseudopotential, density-functional theory foundation, and calculations are shown for systems ranging in size up to 7 nm^3.

1 Introduction

The OCEAN code was introduced several years ago[1–3] as a core-level spectroscopy complement to the valence spectroscopy package AI2NBSE.[4] The purpose of both codes is to provide predictive calculated electronic response, specifically in the near-edge X-ray or optical/UV energy ranges, respectively. Now, the valence and core-level codes are combined into a single package which retains the name OCEAN.[5] In this paper I will detail a number of advances that have been made that improve the accuracy and efficiency of OCEAN as well as extending the scope of materials and spectroscopies that can be simulated. This paper is not intended as a comprehensive review of the underlying theory, but it covers details of the implementation. In particular, I highlight specific approximations made in OCEAN that are not necessarily universal. The OCEAN code is not the only Bethe–Salpeter equation (BSE) based code. For many years BSE has been used for *ab initio* calculations of both valence[6–8] and near-edge X-ray spectra,[9] and is now incorporated into many codes.[10–13] Nor is BSE the only method for calculated optical or X-ray spectra. A large number of computational tools based on a variety of theoretical methods are available.[14,15]

Before continuing, it is instructive to clarify what is meant when OCEAN is referred to as predictive. I contrast this with approaches that are descriptive, and the important distinction is that of the method's primary purpose. In a descriptive approach a parametrized model, cleverly constructed to capture the relevant effects and excitations, is fit to experimental data. From this fit one infers relative strengths and weights of modeled effects in the system being studied. The OCEAN code is predictive in the sense that the system-dependent adjustable parameters are limited to the structural description. The structure determines the species of atoms and their locations. Parameters that are applicable to a wide variety of systems, such as the choice of density functional (see Section 2.1.1), or to a specific atomic species, such as the core–hole lifetime (see Section 2.1.2), are considered acceptable. Removing these parameters (the universal density functional or reliable decay lifetime calculations) are, of course, goals in the condensed matter community, but will be left to others.

To begin I must clearly define the scope of the problem. As stated above, I am interested in calculating the electronic response to perturbations that range in energy from optical to X-ray. More narrowly, the electronic response of condensed systems, liquids or solids (though gas phase or molecular systems are possible as well). The temperature can range from cold to several times room temperature, subject to the constraint that the temperature be small compared to the Fermi level (metals) or bandgap (insulators), $k_\beta T \ll E_\mathrm{F}, E_\mathrm{g}$. This allows the temperature to be treated as a perturbation on top of a 0 K calculation. A pressure range from 0 Pa. to many GPa. is acceptable. These constraints allow for a wide range of possible systems and conditions from battery cathodes to water splitting catalysts to minerals within the Earth's mantle.

The calculation of a material's spectrum involves first the calculation of the ground-state electronic structure of that material. This can be thought of hierarchically. The physical structure of a material, its atoms and their coordinates, determines the electronic structure, and, in turn, the electronic response or spectra. The goal is to provide an overview of how the electronic structure and the electronic response are treated within the OCEAN code.

The format of this paper will continue as follows. After a brief comment on the structure of the code, I give a broad overview of the theoretical underpinnings of OCEAN as well as

---

Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA. E-mail: john.vinson@nist.gov

some implementation details specific to OCEAN and the types of calculations OCEAN is capable of. Then I discuss a number of improvements and advances that have been made to the code before finally presenting a number of examples that showcase the breadth and utility of OCEAN.

1.1 Structure of the OCEAN code

The OCEAN package is a collection of scripts and executable programs. Density-functional theory (DFT) calculations are a vital part of OCEAN, but OCEAN is reliant on external codes for these calculations. After parsing the input file and carrying out initial parameter checks, the OCEAN package runs through several stages. Each stage is controlled by its own script, which allows each stage to be run individually. First is the "OPF" (optimal projector function) stage which calculates details of core-level orbitals (see Section 2.1.2). Next, the "DFT" stage carries out DFT calculations (see Section 2.1.1) to generate electron orbitals which are used in the the BSE calculations. After this is the "PREP" stage which serves as an intermediate step to allow OCEAN to interface with different external DFT codes that might have different output file formats. Next comes the "SCREEN" stage which calculates the dielectric response which screens the electron–hole attraction in the BSE. Finally, the "BSE" stage generates spectra using the BSE.

1.2 Measurements and convergence

With OCEAN one is able to make "computational measurements," that is, probe the theoretical assumptions that are in the code as well as the structural assumptions that are used as input. Unlike reality, where the atomic movement in a system is the dependent variable, changing with applied external forces or conditions, in simulations the atomic positions are the independent (and only free) parameter. This means that one is able to measure how structural changes affect spectra. Additionally, by changing aspects of the underlying theoretical approximations one can observe the effect of that approximation on the electronic structure and response. For instance, the exchange–correlation potentials of DFT are often judged by structural properties, cell volume or bulk modulus, whereas core-level spectroscopy is sensitive to the local shape of the electron orbitals. Computational spectroscopy provides a measurement of the underlying electronic structure beyond energy levels or the density response.

The OCEAN code relies on both uncontrolled approximations, such as the use of the Bethe–Salpeter equation to describe excited states or DFT to describe the ground state, and controlled approximations, such as the use of finite sums instead of integrals over electronic states. It is important that calculations are converged with respect to all controllable approximations, irrespective of wether this convergence improves agreement with measured data.

2 Theory overview

2.1 Ground state

Before simulating the electronic response of a system, one must first simulate the system's ground state. This is because the terms needed to solve the response are built diagrammatically out of the ground-state electron wave functions. While a Hamiltonian can be written down for condensed matter systems with only very minor approximations, such a complete description is not tractable. To proceed a number of approximations are made that substantially reduce the size and complexity of the Hamiltonian.

First, the nuclei and electrons are decoupled, writing the many-body wave function as

$$|\xi(\vec{\mathbf{R}})\rangle|\Psi(\vec{\mathbf{r}};\vec{\mathbf{R}})\rangle \tag{1}$$

where the electron's wave functions Ψ depend explicitly on the electron coordinates $\vec{\mathbf{r}}$, but parametrically on the atomic coordinates $\vec{\mathbf{R}}$. This is the Born–Oppenheimer approximation. It allows one to treat the nuclei as fixed classical potentials and separately solve the electronic Hamiltonian. Often an argument is made that the nuclei move slowly compared to the timescale of the electronic excitations. Importantly, this approximation treats the vibrational excitations as having negligible energy.[16] In what follows spectra are assumed to depend only parametrically on atomic positions, and the atomic positions remain fixed during the excitation.

Next, the material is treated with periodic boundary conditions. This means that the electron orbitals ψ can be written in terms of bands n, their spin σ, and their crystal momenta $\mathbf{k}$ within the Brillouin zone,

$$\psi_{nk\sigma}(r) = \mathrm{e}^{ik\cdot r}u_{nk\sigma}(r) = \mathrm{e}^{ik\cdot r}\sum_{G}\mathrm{e}^{iG\cdot r}C^{G}_{nk\sigma} \tag{2}$$

in either real-space or as a sum over Fourier coefficients C^G, where G are reciprocal lattice vectors. The number of Fourier coefficients is limited by defining a plane-wave cut-off energy $E_\mathrm{c} \leq G^2/2$. Despite the use of periodic boundary conditions, non-crystalline systems can still be simulated with the OCEAN code, such as water against a thin membrane[17] or vibrationally disordered systems,[18] by using large cells with hundreds or a few thousand atoms and checking for the effects of artificial periodicity.

2.1.1 DFT. One-electron orbitals are calculated using density–functional theory (DFT).[19] Briefly, the difficult to calculate electron–electron interactions are replaced by a fictitious Kohn–Sham potential V_KS $[n]$,[20] a functional of the electron density n,

$$H[n] = -1/2\nabla^2 + V_\mathrm{ext} + V_\mathrm{H}[n] + V_\mathrm{KS}[n] \tag{3}$$

where the external potential V_ext is the sum of the atomic potentials (and any applied external potential), and the Hartree potential $V_\mathrm{H}[n]$ is also dependent on the electron density. This means that for a given electron density the many-body Hamiltonian is reduced to a one-particle Hamiltonian, whose eigenstates can be solved using traditional linear algebra techniques. The electron density of a material is not known *a priori*, and an iterative solution is used to converge the density such that the resultant electron density from the N-lowest eigenstates of $H[n]$ is the same as the input density.

The OCEAN code relies on external codes to carry out the necessary DFT calculations to generate the electron density $n(r)$

and electron eigenstates $\{\psi_{nk}\}$ and energies $\{\varepsilon_{nk}\}$. At present, OCEAN interfaces with both QUANTUM ESPRESSO[21] and ABINIT,[10] and it is able to construct their input files, run the codes, and gather and parse their outputs. Adding new plane-wave, psuedopotential DFT code interfaces can be accomplished by writing their output files in a format readable by OCEAN or by extending the OCEAN wave function parser. The examples shown here were calculated with QUANTUM ESPRESSO version 6.7 (version 7.0 for water).

2.1.1.1 Exchange–correlation functionals and DFT+U. While DFT is an exact theory, the exact exchange–correlation functional $V_{\text{KS}}[n]$ is not known. The simplest Kohn–Sham exchange–correlation potential is referred to as the local-density approximation (LDA). Within the LDA the value of the Kohn–Sham potential is based on calculations of the homogenous electron gas. More sophisticated exchange–correlation potentials are possible, *e.g.*, taking into account the local gradients of the density as well as is done in the class referred to as generalized gradient approximations (GGA) such as the popular Perdew–Burke–Ernzerhof (PBE) functional.[22] One-step further, *meta*-GGA functionals include terms that depend on the second derivative of the density or the kinetic energy density of the occupied orbitals, like the so-called strongly constrained and appropriately normed (SCAN) functional.[23] A different approach for improving the reliability of functionals is to add in some amount of the true exchange interaction in place of the exchange component of the density functional – referred to as exact exchange (EXX).

Lastly, local and semi-local DFT is known to perform poorly for systems with highly-localized states near the valence-band maximum or conduction-band minimum, *i.e.*, 3d and 4d transition metals. This has lead to the use of a Hubbard-*U* term (DFT+*U*), where an orbital-specific term is added to the Hamiltonian. In its simplest form, this term is an on-site repulsion that penalizes doubly occupying the d-orbitals, stabilizing high-spin configurations and opening a Hubbard gap.

2.1.1.2 GW corrections. Despite the formalism of DFT designed to produce only the correct ground-state density of a system, DFT orbitals have been widely used in spectroscopy. DFT orbitals often suffer from incorrect relative energies, leading to band gaps and band widths that are too small. Using many-body perturbation theory, the DFT orbitals can be treated as quasiparticles and their energies corrected by the self-energy operator Σ. The first-order energy correction is given by

$$\varepsilon_i = \varepsilon_i^{\text{KS}} + \langle \Psi_i | \Sigma(\varepsilon_i) - V_{\text{KS}} | \Psi_i \rangle \tag{4}$$

where the self-energy takes the place of the Kohn–Sham potential. Note that the self-energy depends on the correct quasiparticle energy, and, depending on the approximations used, the self-energy can have an imaginary component that is the inverse of the quasiparticle lifetime. As developed by Hedin, the self-energy operator can be approximated from one-electron orbitals produced either with the a Kohn–Sham potential (DFT) or a bare exchange interaction (Hartree-Fock). The first-order approximation for the self-energy operator is given by $\Sigma = \text{GW}$, leading to its designation as the GW approximation.[24]

OCEAN does not carry out GW calculations, but results from external GW codes can be included in an OCEAN calculation. Often calculations keep the DFT orbitals constant and only energy-corrections to the DFT eigenvalues are calculated, the so-called G^0W^0 approximation. So long as the exchange–correlation functional and pseudopotentials are the same, energy corrections from one code can be applied directly to eigenvalues of another. Energy corrections can be supplied to OCEAN for all states explicitly Σ_{nk}, by band Σ_n, or by energy $\Sigma(\omega)$, and can be either real- or complex-valued. Alternatively, a simple band-gap and stretch can be applied where the valence and conduction band energies are stretched by separate amounts, and the band gap set or changed by a specific value. Quasi-particle self-consistent GW calculations have been used with OCEAN,[25] but require inputting the electron orbitals from the GW calculation and this has not yet been automated.

2.1.2 Core-level orbitals. The use of pseudopotential DFT to supply the valence and conduction band electron orbitals for OCEAN presents a problem for core-level spectroscopy: the absence of core-level orbitals. This shortcoming is solved by using an auxiliary atomic DFT code which is provided as part of OCEAN or, optionally, a modified version of the ONCVPSP code.[26,27]

In order to calculate transitions and interactions between core-level orbitals and orbitals from a pseudopotential-based DFT, OCEAN makes use of optimal projector functions (OPFs), a method based upon the projector augmented wave formalism.[28] Details of calculating the OPFs and recent improvements are discussed elsewhere.[29] The OPFs provide a complete basis set for describing a limited space of DFT orbitals – from the valence bands up to approximately 6 Ryd above the Fermi level and within a small sphere around an atom. The sphere radius is at or slightly larger than the pseudopotential radius, in the range of 1 a.u. to 2.5 a.u. This basis allows the pseudopotential DFT orbitals to be augmented to include the correct all-electron character.

2.1.2.1 Energy alignment. OCEAN and the DFT codes it interfaces with are based on pseudopotentials, and core-levels are ignored. Thus changes in the core-level energy due to chemical surroundings are not available. The energies from the separate, atomic DFT code necessarily cannot account for differences in the electronic structure at the atomic site in a given structure. As an example, an approximately 4 eV energy shift is found experimentally in ammonium nitrate between the N 1s orbitals in the N^{3-} site of ammonium and the N^{5+} site of nitrate.[18]

Energy alignments are calculated by first noting that, by using an external atomic code, the calculation is being done within the frozen core approximation. The core-level orbitals are independent of the material. Therefore, the DFT energy of the core orbital γ can be given by some constant X_γ plus the total external electrostatic potential experienced by the orbital V_{KS}. This total external potential is the total potential of the pseudopotential DFT calculation. Finally the adiabatic relaxation of the system is included in response to removing the core orbital

$$\varepsilon_\gamma = X_\gamma + \int \mathrm{d}\mathbf{r}\rho_\gamma(\mathbf{r})V_{\mathrm{KS}}(\mathbf{r}) - 1/2\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\rho_\gamma(\mathbf{r})W(\mathbf{r},\mathbf{r}')\rho_\gamma(\mathbf{r}'), \quad (5)$$

where $\rho_\gamma = |\psi_\gamma|^2$ is the density of the core orbital, and W is the screened Coulomb operator that is already constructed for use in the BSE. Previously, the extent of the core orbitals was ignored and the potential and screening were evaluated over a delta function at the position of the atom. This method for determining core-level shifts in OCEAN has been used between sites within a cell,[18,30] aligning within and between disordered cells such as in liquids,[31] and between different compounds.[32] Further work is needed to benchmark it against experimental data and other theoretical methods for determining core-level alignment.

2.2 The Bethe–Salpeter equation method for absorption and related spectroscopies

The Bethe–Salpeter equation (BSE) is a series expansion of the two-particle propagator.[33] While originally applied to the ground-state nucleus of a deuteron, it is used widely in condensed matter to describe the behavior of neutral particle–hole excitations. The reader is referred to ref. 34 and 35 and references therein. At present, OCEAN makes use of the Tamm-Dancoff approximation.[36] Several additional approximations are required in order to determine the electronic response of a system using the BSE.

The first set of approximations is that the electrons are treated either non-relativistically or, at most, in a scalar relativistic manner. The exchange of real or virtual photons is done without accounting for the finite speed of light. Electrons have two-component, up or down spins, not 4-component Dirac spinors. The magnitude of the spin–orbit splitting of core-level electrons is calculated in a scalar relativistic framework and applied perturbatively in the angular momentum and spin, l, m_l, σ, basis, *e.g.*, between the $2p_{3/2}$ L_3 and $2p_{1/2}$ L_2. The spin–orbit splitting is applied as if it were an interaction term $\hat{\xi}$ in the BSE Hamiltonian.

Importantly, in the BSE only two explicit excited quasiparticles exist at a time – an excited electron and a hole. There is an implicit shift in the crystal momentum between electron and hole that conserves the photon's momentum for absorption or emission. The effective BSE Hamiltonian used by OCEAN can be written in the basis of a pair of occupied (a, b) and unoccupied (i, j) states

$$H^{\mathrm{BSE}}_{ia,jb} = (\varepsilon_i - \varepsilon_a)\delta_{ij}\delta_{ab} + \hat{\xi}_{ab}\delta_{ij} - W_{ia;jb} + X_{ia,jb} \quad (6)$$

which includes the bare energies ε, the pseudo-interaction of the spin–orbit ξ, and the two electron–hole interaction terms, the direct W and exchange X. The quasiparticle energies ε can be complex-valued, with an imaginary component inversely proportional to the lifetime of the quasi-electron or quasi-hole single-particle excitation. The unoccupied states $\{i, j\}$ are indexed by band, k-point, and spin, and the same is true for occupied states when the hole is in the valence bands. For core-hole spectroscopy, the occupied states $\{a, b\}$ are deep core levels that are indexed by azimuthal quantum number m and spin. The core level also is identified by atomic site, principle quantum number n and angular momentum l.

The BSE is evaluated using a finite k-point mesh of electron and hole states in place of an integral over the Brillouin zone. The mesh subdivides the reciprocal lattice vectors, and the number of k-points in each direction should be proportional to the length of that reciprocal lattice vector. Note, that the lengths of the reciprocal lattice vectors depend on both the lengths of the lattice vectors and the angles between them. The number of conduction bands included in the calculation is also limited. This has the effect of limiting the spectral range of the calculation, but, because the DFT single-particle states are not eigenstates of the BSE, convergence of even the low energy part of a calculated spectrum must be checked with respect to the number of conduction bands. As the occupied states at a given k-point are finite they can all be included in the BSE calculation, but this is not done. The core levels are energetically separated enough to be neglected for valence calculations and *vice versa*, while the core orbitals between atoms have no overlap and are treated independently for core-level spectroscopy.

2.2.1 Electron–hole Interactions. In addition to the non-interacting electron and hole energies, two electron–hole interaction terms are included in the BSE. The direct interaction is responsible for the excitonic binding between electron and hole, and is screened by the dielectric response of the electrons. The matrix elements of the direct are taken to be

$$\langle i,a|W|j,b\rangle = \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\psi_i^*(\mathbf{r})\psi_a(\mathbf{r}')W(\mathbf{r},\mathbf{r}')\psi_j(\mathbf{r})\psi_b^*(\mathbf{r}')\delta_{\sigma_a\sigma_b}\delta_{\sigma_i\sigma_j} \quad (7)$$

with electron states i, j and hole states a, b. Here $W = \varepsilon_\infty{}^{-1}v$ is the Coulomb interaction screened by the static dielectric response of the electrons in the system. (It is assumed that the ions do not respond on a fast enough time scale to contribute.) Extensions to dynamic screening are possible,[37] but enter in with a resonant energy denominator which is peaked at the energy difference between energy of the BSE eigenstate and the constituent single particles i and a. Since the excitonic binding is typically small compared to the band gap, the use of only the static screening W ($\omega = 0$) is justified.

The exchange interaction involves the electron–hole pair combining, emitting a virtual photon, and then reappearing, and it involves the unscreened Coulomb operator

$$\langle i,a|X|j,b\rangle = \int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\psi_i^*(\mathbf{r})\psi_a(\mathbf{r})v(\mathbf{r},\mathbf{r}')\psi_j(\mathbf{r}')\psi_b^*(\mathbf{r}')\delta_{\sigma_i\sigma_a}\delta_{\sigma_j\sigma_b} \quad (8)$$

Note the spatial indices and spin selection rules are changed from the direct.

2.2.1.1 Valence-level interactions. For solving the BSE with a valence–hole–conduction–electron excitation, *e.g.*, optical/UV excitations or the final state of RIXS (see Section 2.3.5), the OCEAN code closely follows earlier work,[8,38] and I will only briefly summarize it here. The electron and hole states are down-sampled from the Fourier coefficients **G** from the DFT code onto a real-space mesh **x**. The mesh is aligned with the lattice vectors with a typical spacing of 0.5 a.u. to 1.0 a.u., or a plane-wave energy cut-off range 10 Ryd to 40 Ryd. The direct interaction is evaluated in real-space with a Fourier transform of the *k*-point mesh defining a supercell in which the exciton is confined. Evaluation of the direct is the most time-consuming part of the valence BSE as both the electron and hole components must be evaluated on the full *x*-mesh. The direct interaction is evaluated in reciprocal space where it is diagonal.

2.2.1.2 Core-level interactions and the local basis. For the core-level BSE, both the direct and the exchange interaction are evaluated in real-space. Exploiting the locality of the core-level orbital, both the direct and exchange interactions are written as partial wave expansions around the atomic site of interest. It can be immediately seen from eqn (8) that the exchange interaction is limited to the region immediately around the atom because the core orbital must be present at both spatial coordinates. The direct interaction does not have this limitation, and the l = 0 component is very long ranged. Only even powers of l enter into the direct interaction, and by the l = 2 term the interaction decays quickly enough that a local treatment is sufficient. The l = 0 component is broken into two pieces

$$\begin{aligned} W^{(0)}(r) &= \left[W^{(0)}(r) - W^{(0)}(r_{\mathrm{c}})\right]\Theta(r_{\mathrm{c}} - r) \\ &\quad + \left[W^{(0)}(r_{\mathrm{c}})\Theta(r_{\mathrm{c}} - r) + W^{(0)}(r)\Theta(r - r_{\mathrm{c}})\right] \end{aligned} \tag{9}$$

where the first piece is non-zero only within some cut-off radius r_{c}, and the core–hole density has been integrated out to give the core–hole potential W as just a function of the electron coordinate r. The addition and subtraction of the value of W at the cut-off radius ensures that both pieces are smooth.

The exchange and short-range components of the direct are evaluated using the OPFs (Section 2.1.2). Between 2 and 5 projectors are constructed per angular momentum channel, and angular momenta up to l = 3 are included, providing a compact basis of no more than 80 OPFs. The long-range component of the direct interaction is evaluated on the *x*-mesh. It is the long-range component that dominates computational time for large unit cells. The strength of the exchange and short-range, $l \geq 2$, components of the direct are typically reduced to 80% of their calculated values.[39]

2.2.1.3 Spin–orbit splitting. While not an electron–hole interaction, the effects of spin–orbit splitting are treated in OCEAN as part of the interaction Hamiltonian. An electron moving in the electric field of the nucleus will experience an effective magnetic field which, in turn, interacts with the electron's magnetic moment. This is the spin–orbit interaction and takes the form $\hat{\xi} \propto (\mathbf{L}{\cdot}\mathbf{S})$, where $\mathbf{L}$ and $\mathbf{S}$ are the angular momentum and spin operators respectively. The spin–orbit interaction is not diagonal in the L_z, S_z representation, and the usual practice is to use the total angular momentum basis J^2, J_z. The operators L^2 and S^2 are diagonal in both representations. We take the spin–orbit splitting within first order perturbation theory,

$$\Delta_{nlj} = \xi_{nl}\frac{1}{2}\begin{bmatrix} l \\ -(l+1) \end{bmatrix} \quad \begin{matrix} j = l + 1/2 \\ j = l - 1/2 \end{matrix} \tag{10}$$

$$\xi_{nl} = \frac{\alpha^2}{2}\left\langle \phi_{nl}\left|\frac{1}{r}\frac{\mathrm{d}V_{\mathrm{c}}}{\mathrm{d}r}\right|\phi_{nl}\right\rangle \tag{11}$$

where ξ_{nl} is the strength of the spin–orbit splitting and depends on the central potential V_{c} and the electronic orbitals ϕ_{nl}. The calculation of ξ_{nl} is done in the auxiliary atomic code, and is therefore representative of an isolated neutral atom in a low-spin configuration.

The spin–orbit interaction is not diagonal in the basis vectors used in OCEAN. (The direct and exchange interactions as well as the valence and conduction band states all make use of S_z, and the exchange and local part of the direct interaction make use of L_z.) In the L_z, S_z basis, the dot product is expanded $\mathbf{L}{\cdot}\mathbf{S} = 1/2(J^2 - L^2 - S^2)$

$$\hat{\xi}_{abij} = \frac{\xi_{nl}}{2}\langle lm_l, 1/2\ s|J^2 - L^2 - S^2|l'm_l', 1/2\ s'\rangle\delta_{\sigma\sigma'}\delta_{\mathbf{kk}'}\delta_{mm'} \tag{12}$$

where the core level $\{a, b\}$ has angular quantum number l, magnetic quantum number m_l, and spin s. The spin–orbit operator does not mix the conduction bands $\{ij\}$ denoted by k-vector $\mathbf{k}$, band index m and spin σ. Implicitly, the spin–orbit interaction is limited to the core levels of a single atom sharing both principle and angular quantum numbers.

The spin–orbit splitting also leads to a j-dependent lifetime broadening. A hole in the more-bound $2p_{1/2}$ state can decay into a less-bound $2p_{3/2}$ state, but not the converse. The core–hole lifetime broadening values Γ_{nlj} are not calculated by OCEAN. However, experimentally they are seen to be largely independent of the local electronic structure and fairly constant across materials (for a given element and core level). Therefore, they are available in OCEAN as a lookup table based on experimental measurements.[40] The state-dependent broadening is included in the spin–orbit Hamiltonian above by dividing the broadening into j-dependent Γ_{nl} and j-independent $\bar{\Gamma}_{nl}$ components:

$$\Gamma_{nl} = \frac{2}{2l+1}(\Gamma_{nl+} - \Gamma_{nl-}) \tag{13}$$

$$\bar{\Gamma}_{nl} = \frac{l+1}{2l+1}\Gamma_{nl+} + \frac{l}{2l+1}\Gamma_{nl-} \tag{14}$$

where Γ_{nl+} (Γ_{nl-}) is the broadening of the $j = l + 1/2$ ($j = l - 1/2$) level. Physically Γ_{nl-} is the larger of the two, and therefore Γ_{nl} is negative.

This spin–orbit dependent lifetime broadening is added to the real-valued, spin–orbit splitting interaction giving

$$\hat{\xi}_{abij} = \left[(\mathbf{L}\cdot\mathbf{S})(\xi_{nl} - \mathrm{i}\Gamma_{nl}) - \mathrm{i}\bar{\Gamma}_{nl}\delta_{ss'}\delta_{m_l m_l'}\right]\delta_{l,l'}\delta_{\sigma\sigma'}\delta_{\mathbf{kk}'}\delta_{mm'} \tag{15}$$

The negative sign follows the convention that the imaginary part of a self-energy or lifetime broadening is negative for states below the Fermi level. The spin–orbit component of the overall Hamiltonian enters with a minus sign such that the imaginary component of any eigenenergy of the total Hamiltonian will be positive. As in the case of complex GW eigenvalues, the addition of any state-dependent broadening results in a non-Hermitian BSE Hamiltonian.

2.3 Calculating spectra

Within OCEAN the BSE Hamiltonian is never explicitly constructed and is not diagonalized. The code calculates the action of the Hamiltonian on a vector in the electron–hole basis. Each of the aforementioned interaction terms is evaluated in a favorable basis: real-space mesh, reciprocal-space mesh, or local basis. Two iterative solvers are available that are capable of calculating the spectra, while the second is also used to generate the exciton wave functions.

The macroscopic dielectric function is given by

$$\begin{aligned}\varepsilon(\mathbf{q},\omega) &= 1-\frac{4\pi}{\Omega_V q^2}\langle\Phi_0|T^{\dagger}_{\mathbf{q},\hat{\varepsilon}}G_{\mathrm{BSE}}(\omega)T_{\mathbf{q},\hat{\varepsilon}}|\Phi_0\rangle \\ &= 1-\frac{4\pi}{\Omega_V q^2}\frac{1}{N_k}\sum_{ab}\langle\Phi_0|T^{\dagger}_{\mathbf{q},\hat{\varepsilon}}|a\rangle\langle a| \\ &\quad\times\frac{1}{\omega-H_{\mathrm{BSE}}+\mathrm{i}\eta}|b\rangle\langle b|T_{\mathbf{q},\hat{\varepsilon}}|\Phi_0\rangle\end{aligned} \tag{16}$$

where $a, b = \mathbf{k}, c, v$ (or $v \to \gamma$ for core), Ω_V is the unit cell volume, and N_k are the number of k-points included. Because each core level and site τ is treated independently, the complete full-frequency dielectric response would be given by

$$\varepsilon(\mathbf{q},\omega) = 1+[\varepsilon_{\mathrm{val}}(\mathbf{q},\omega)-1]+\sum_{\tau,nl}\left[\varepsilon_{\tau,nl}(\mathbf{q},\omega)-1\right] \tag{17}$$

where $\varepsilon_{\mathrm{val}}$ is the valence-only response and each $\varepsilon_{\tau,nl}$ is the response of a specific n and l of a single atom τ.

2.3.1 Electron–photon interactions. The interaction in terms of the photon vector potential $\mathbf{A}$ within the Coulomb gauge ($\nabla\cdot\mathbf{A} = 0$) are given by

$$H_{\mathrm{int}} = -\mathbf{p}\cdot\mathbf{A}-\mu\cdot(\nabla\times\mathbf{A})+\frac{\alpha^2}{2}\mathbf{A}^2 \tag{18}$$

where $\mathbf{p}$ is the momentum of the electron and $\mu \approx \alpha\mathbf{s}$ is the electron's magnetic moment or approximately the fine structure constant times its spin. The problem is divided into two classes, scattering $\mathbf{A}^2$ or absorption or emission $\mathbf{A}$.

2.3.1.1 Valence. In the case of valence band to conduction band excitations, the matrix elements are proportional to

$$M_{vc\mathbf{k};\mathbf{q}} = \langle c\mathbf{k}|\mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}}|v\mathbf{k}-\mathbf{q}\rangle = \langle v\mathbf{k}-\mathbf{q}; c\mathbf{k}|\mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}}|\boldsymbol{\Phi}_0\rangle \tag{19}$$

where $\boldsymbol{\Phi}_0$ is the ground state, v the valence band index, c the conduction band index, and $\mathbf{q}$ is the photon momentum. By using an explicit shift between the conduction states at $\mathbf{k}$ and valence states at $\mathbf{k}-\mathbf{q}$ these matrix elements are a simple dot product of the Bloch functions. For optical excitations, the magnitude of $\mathbf{q}$ is small. The same transition operator can also be used to simulate valence band electron energy loss (EELS) which can be measured for a range of momenta.

2.3.1.2 Core. For core-level excitations the choice of electron-photon interaction is more varied. In all cases the matrix elements can be written

$$M_{\gamma c\mathbf{k};\mathbf{q}\varepsilon} = \sum_{\nu,l,m}\langle c\mathbf{k}|\nu lm\rangle\langle\nu lm|T_{\mathbf{q},\hat{\varepsilon}}|\gamma\rangle = \langle\gamma; c\mathbf{k}|T_{\mathbf{q},\varepsilon}|\Phi_0\rangle \tag{20}$$

where the core level is given by γ, the photon operator T can depend on both momentum $\mathbf{q}$ and polarization $\hat{\varepsilon}$, and the OPF basis $\{\nu lm\}$ is used to calculate the transitions.

For scattering, the relevant class of experiments are non-resonant inelastic X-ray scattering (NRIXS) also called X-ray Raman scattering (XRS). Dropping the polarization, this operator is the same as is used to simulate EELS.

$$T_s = \vec{\varepsilon}_1\cdot\vec{\varepsilon}_2\mathrm{e}^{i\mathbf{k}_1\cdot\mathbf{r}}\mathrm{e}^{-i\mathbf{k}_2\cdot\mathbf{r}} \propto \mathrm{e}^{i\mathbf{q}\cdot\mathbf{r}}$$

The momentum transfer $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_2$ is the momentum that the scattered photon imparts to the electronic system. For small momenta this operator resembles a dipole, but in general can include much higher orders. The matrix elements are computed using a localized basis set for the conduction-band electrons $|\nu lm\rangle = R_{\nu l}(r)Y_{lm}(\hat{r})$. The transition operator can be expanded into spherical Bessel functions j_l

$$\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi\sum_{l=0}\sum_{m=-l}^{l} i^l j_l(kr)Y_{lm}(\hat{k})Y^*_{lm}(\hat{r}) \tag{21}$$

The sum over angular momenta is truncated at the sum of the core level and the largest local basis momenta, typically no more than $l = 3$ by the selection rules governing integration over three spherical harmonics, *i.e.*, Gaunt coefficients. In practice, the matrix elements are evaluated by real-space integration over both the radial and angular parts

$$\begin{aligned}\langle nlm|\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}|\nu'l'm'\rangle &= 4\pi i^l\int r^2\mathrm{d}rR_{nl}(r)j_l(kr)R_{\nu'l'}(r) \\ &\quad\times\int\mathrm{d}\hat{r}Y^*_{lm}(\hat{r})Y_{l'm'}(\hat{r})P_l(\hat{k}\cdot\hat{r})\end{aligned} \tag{22}$$

An early version of OCEAN used a power expansion of the operator, but this is only convergent for strongly localized core holes (significantly smaller than one Bohr).

For absorption and emission the matrix elements are also calculated in real space using the localized basis. Interactions are calculated to at most quadrupole order, which includes a single $\mathbf{k}\cdot\mathbf{r}$ term.

$$T_a = \vec{\varepsilon}\cdot\mathbf{r} - \mathrm{i}/2(\vec{\varepsilon}\cdot\mathbf{r})(\mathbf{k}\cdot\mathbf{r}) \tag{23}$$

In the case of linearly polarized light, this can be calculated in the same manner as the scattering matrix elements, $\langle nlm|T_a|\nu'l'm'\rangle$.

2.3.2 Haydock recursion. The primary method for calculating spectra with OCEAN is using the Haydock recursion method[41] which relies on a symmetric Lanczos tridiagonalization.[42] This method was used by Benedict and Shirley,[38] and is included in other BSE and non-BSE spectroscopy codes.[11,43] At present, only a symmetric, Hermitian Lanczos is implemented, and therefore the Haydock method cannot be used when the electron or hole quasiparticles have non-uniform complex eigenvalues (GW corrections or $L_{2,3}$ broadening). A uniform, state-independent lifetime broadening (η in eqn (16)) is always used with the Haydock method. This broadening is typically set to match the expected broadening from the core–hole lifetime.

The advantages of using the Haydock recursion method are two-fold. One, the number of iterations required to produce a spectrum (rank of the tridiagonal matrix) is not dependent on the rank of the BSE Hamiltonian. It depends directly on the spectral range included in the BSE (the energy range of the conduction bands) and inversely on the resolution (the core–hole lifetime broadening). While the BSE Hamiltonian can easily be of rank 10 000, the number of iterations needed to converge the Haydock is typically around 100. Second, though general to any iterative solve, as noted above the BSE Hamiltonian is never explicitly constructed, reducing computational time and memory requirements.

2.3.3 GMRES. The second method for calculating the BSE implemented within OCEAN is the generalized minimal residual (GMRES) method.[44] The GMRES method is suitable for iteratively solving matrix equations of the form $Ax = b$. Defining the exciton vector $x(\omega)$,

$$|x_{\mathbf{q},\hat{\varepsilon}}(\omega)\rangle = \frac{1}{\omega - H_{\mathrm{BSE}} + \mathrm{i}\eta} T_{\mathbf{q},\hat{\varepsilon}}|\Phi_0\rangle \tag{24}$$

GMRES is used to solve

$$(\omega - H_{\mathrm{BSE}} + i\eta)|x_{q,\hat{\varepsilon}}(\omega)\rangle = T_{q,\hat{\varepsilon}}|\Phi_0\rangle \tag{25}$$

and, substituting into eqn (16), the dielectric function is

$$\varepsilon(\mathbf{q},\omega) = 1 - \frac{4\pi}{\Omega_V q^2}\langle\Phi_0|T^{\dagger}_{\mathbf{q},\hat{\varepsilon}}|x_{\mathbf{q},\hat{\varepsilon}}(\omega)\rangle \tag{26}$$

Note that the exciton vector $x(\omega)$ is not the same as an eigenstate of the BSE Hamiltonian. Rather, it is a projection of all the eigenstates accessible *via* the electron-photon operator T acting on the ground state Φ_0 and within an energy range set by the resonant energy denominator with width η. Similar to the Haydock method, η is a state-independent broadening parameter, typically the core–hole lifetime broadening, but additional, state-dependent broadening can also be included within the BSE Hamiltonian.

In OCEAN, right preconditioning with a diagonal preconditioner is used,

$$M_{ij}{}^{-1}[\omega,\zeta] = \delta_{ij}\frac{\omega - H_{jj} - i\zeta}{(\omega - H_{jj})^2 + \zeta^2}, \tag{27}$$

which takes as parameters the energy being solved for ω and a broadening term ζ, which is typically set to 13.6 eV. In the case of no electron–hole interactions (and $\zeta = 0$) this preconditioner is exact. The broadening term ζ must be sufficient to account for the interactions, primarily the direct, as well as any spin–orbit splitting. A more sophisticated preconditioner that accounted for the block-diagonal mixing of the spin–orbit interaction might be more efficient for transition metal $L_{2,3}$ edges, but it has not been attempted.

Typically, approximately 100 iterations are required to converge the vector $x(\omega)$ for a single energy, making the GMRES method roughly N_ω times more computationally expensive than the Haydock recursion. The GMRES method is restarted every 80 iterations to avoid the growing costs of a large Krylov space. If the energy points are sufficiently close compared to the core–hole broadening, $|\omega_i - \omega_{i-1}| < 3\eta$, then the previous vector x (ω_{i-1}) is used as an initial guess for $x(\omega)$. In all other cases the initial guess is set to 0. Additionally, when the energy points are close compared to the preconditioner broadening $\omega_i - \omega_{i-1} < \zeta/4$ the Krylov space is recycled from the previous iteration and the preconditioner is held centered at the previous energy until a restart. Together these two modifications can reduce the number of iterations required for repeated energy points significantly.

2.3.4 X-ray emission. The calculation of X-ray emission spectra (XES) is carried out in OCEAN at the DFT level. It is equivalent to a projected density of states calculation, where the projector is given by the photon operator and the core orbital

$$\sigma_{\mathrm{XES}}(\omega) \propto \sum_{nk} |\langle\psi_{nk}|T|\gamma\rangle|^2 \delta(\omega - \varepsilon_{nk} - \varepsilon_\gamma) \tag{28}$$

This is calculated in OCEAN using the same Haydock approach as absorption. There are no electron–hole interactions because both the initial and final states contain only a single hole and no excited electron. This approach neglects any relaxation of the local valence states in response to the core hole. Recent work has been done using the BSE to calculate emission by considering large systems with a single core–hole as the initial state, *i.e.*, a defect calculation, and then calculating the BSE response of creating an excited electron in the core level and valence hole.[45]

2.3.5 Resonant inelastic X-ray scattering. Valence resonant inelastic X-ray scattering (RIXS) is an increasingly widely used technique for probing low-energy excitations in a system using X-ray in/X-ray out spectroscopy. In RIXS, the system is probed by measuring the momentum transfer and energy loss between the out-going and incoming X-rays. These low-energy excitations can be any excitations in the system that might couple to the core-level exciton created by the initial X-ray absorption event, *e.g.*, phonons, magnons, *etc.*, but within OCEAN only electron–hole excitations can be calculated.

RIXS spectra are calculated by first solving for the initial X-ray absorption exciton vector using GMRES.

$$\begin{aligned}\sigma_{\mathrm{RIXS}}(\mathbf{q},\varepsilon,\omega;\mathbf{q}',\varepsilon',\omega') \propto &\langle\Phi_0|T^{\dagger}_{\mathbf{q},\hat{\varepsilon}}G(\omega)T_{\mathbf{q}',\hat{\varepsilon}'}G(\omega-\omega') \\ &\times\ T^{\dagger}_{\mathbf{q}',\hat{\varepsilon}'}G(\omega)T_{\mathbf{q},\hat{\varepsilon}}|\Phi_0\rangle\end{aligned} \tag{29}$$

$$\propto \langle x_{\mathbf{q},\hat{\varepsilon}}(\omega)|T_{\mathbf{q}',\hat{\varepsilon}'}G(\omega-\omega')T^{\dagger}_{\mathbf{q}',\hat{\varepsilon}'}|x_{\mathbf{q},\hat{\varepsilon}}(\omega)\rangle \tag{30}$$

The emission spectra for each incoming X-ray energy ω are generated using the Haydock method and valence BSE as a function of the energy loss $\omega - \omega'$. As for valence calculations, RIXS are calculated using a finite momentum $\mathbf{q}$ which can be set to explore the q-dependence of the RIXS spectra. For a large number of incoming X-ray energies, often referred to as RIXS maps, the repeated use of the GMRES and Haydock iterative methods for each energy may become less efficient than solving the two, core and valence, BSE eigensystems and building the spectra from sums over eigenstates.[46]

RIXS at spin–orbit-split edges, *e.g.*, $L_{2,3}$ or $M_{4,5}$, present the additional complication that spin must be included in the valence-level calculation. This is true even if the valence and conduction bands of the system being studied are degenerate with respect to spin. The core-level spin–orbit splitting mixes the spin of the electron and hole components of the exciton. Previous versions of OCEAN were limited to singlet excitations in the valence solver, but this has now been alleviated, allowing RIXS calculations for spin–orbit-split edges. The computational cost is increased by a factor of four over what is needed for a singlet calculation if the valence and conduction bands are spin degenerate (factor of two if they are not). At present this functionality is only implemented for RIXS calculations.

As an example, RIXS calculations have been carried out for the sulfur $L_{2,3}$ edge of CdS. The energy of the occupied Cd 4d-orbitals was adjusted by applying a Hubbard-U correction of 4.2 eV.[48] No valence- or conduction-band spin–orbit was included in the DFT calculation. A uniform broadening of 0.1 eV was used to calculate both the XAS and RIXS. The spin–orbit splitting in sulfur is quite small, around 1 eV, but it is still sufficient to mix the spins. In Fig. 1, the calculated X-ray absorption spectrum (XAS) is shown compared to the experiment from ref. 47. Also shown, is the ratio of spin-aligned electron and hole to the total exciton (both spin-aligned and anti-aligned). It is clear that both singlet and triplet interactions are important for the final-state valence-level BSE when carrying out RIXS at the $L_{2,3}$ edge.

The calculated CdS S $L_{2,3}$ edge RIXS map is shown in Fig. 2(a) while a comparison to experiment is shown in Fig. 2(b). The CdS RIXS is dominated by the very flat S 3s (emission around 146 eV) and Cd 4d bands (emission around 150 eV), which, other than changes in intensity, show almost no dependence on the incident photon energy. The emission from the upper valence bands (emission above 153 eV) show only slight changes in spectral shape with excitation energy. The overall lack of strong excitonic effects in the measured emission explains the success of previous DFT-based RIXS calculations.[47,49] Additionally, the peaks from the sulfur 3s bands are much too narrow in the calculation, resulting in an exaggerated peak intensity. The agreement between the calculation and measurement is good. However, as is typical, the calculated RIXS maps shows more variation with excitation energy than the measured map. This is particularly evident in the upper valence bands, with emission energies above 153 eV. Some of the apparent lack of variation in the experiment is due to incoherent emission. Phonon scattering off of the core hole during the RIXS process has the effect of breaking the momentum dependence between the core hole and excited electron. This in turn allows the final valence hole to occupy any part of the Brillouin zone as in non-resonant X-ray emission. OCEAN is not able to calculate this phonon coupling, and this effect is not accounted for in the calculated spectra.

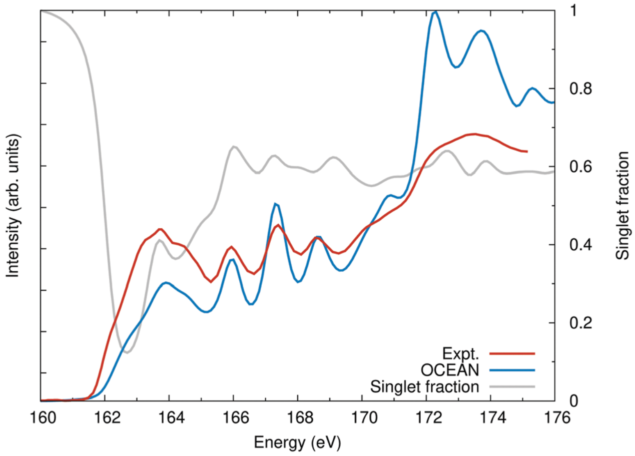

Fig. 1 The calculated sulfur $L_{2,3}$ edge of CdS compared with experiment.[47] Also shown, the singlet fraction is the ratio of how much of the exciton is due to a spin-up hole and spin-up electron or a spin-down hole and a spin-down electron as a fraction of the total excitonic weight. Without spin–orbit, linearly polarized X-rays will only populate spin-singlet excitons.

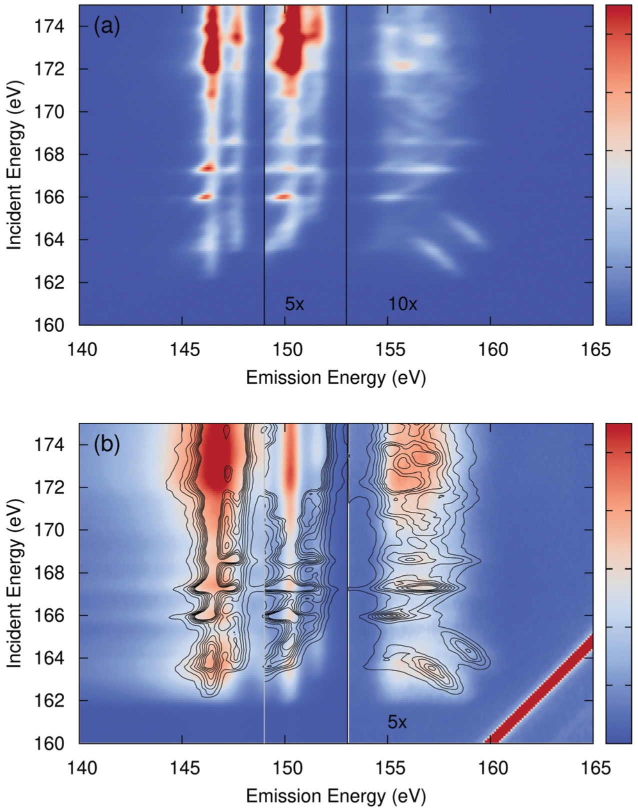

Fig. 2 (a) The calculated sulfur $L_{2,3}$ edge RIXS of CdS as a function of incident and emitted photon energy (intensity in arbitrary units). The region with emission between 149 eV and 153 eV, emission from bands primarily associated with the Cd 4d orbitals, has had the intensity increased by 5×. The region above 153 eV has been increased by 10×. (b) the measured spectra from ref. 47 with contour lines from the calculation overlaid. Above an emission energy of 153 eV the experiment has been increased by 5×.

3 Improvements and developments to OCEAN

In this section I briefly summarize the additions to the OCEAN code over the past 7 years since the previous major version and corresponding publication.[2] These improvements cover enhancements to usability, speed, accuracy, and functionality.

3.1 Input file and defaults

The input file format has been substantially updated, but OCEAN is still capable of parsing and using an input file from version 2. Input flags have been grouped into hierarchies, and both plain text and JavaScript Object Notation (JSON) formats are supported. During setup, OCEAN automatically writes a single file that contains all of the relevant information needed to recreate a calculation, including both the user-supplied information as well as default values for any optional inputs.

The number of required inputs has been reduced as much as possible, as can be seen in Fig. 3. A calculation can be carried out specifying only the DFT program, type of calculation (*e.g.* XAS) and edge (*e.g.* Ti 2p), lattice parameters, and atomic coordinates. The *k*-point mesh for both the BSE final states and screening orbitals is determined by noting that each dimension of the *k*-point mesh subdivides one of the reciprocal lattice vectors. By default, the spacing of this division is set to be no more than 0.33 a.u.^{-1} for the BSE and 0.39 a.u.^{-1} for the screening. This is sufficient to converge the screening calculation in typical systems and provides a reasonable starting point for the BSE. The number of conduction bands included is approximately enough for 50 eV above the Fermi level for the BSE and 100 eV for the screening. These are calculated based on the energy levels of a particle in a box with same volume as the unit cell. The real-space sampling of the orbitals for the BSE, used for all the interactions in a valence-level calculation or the long-range interactions for the core (see Section 2.2.1), is set to divide the lattice vectors by no more than 1 a.u. The other input parameters are not dependent on the unit cell volume.

3.1.1 Pseudopotential support and database. Pseudopotentials are required for carrying out the DFT calculations to generate the electron density and orbitals needed in OCEAN. Additionally, for core-level spectroscopy, OCEAN requires the pseudopotentials for generating the OPFs (see Sections 2.1.2 and 2.2.1.2). Unlike calculations of a material's structure, BSE calculations include large numbers of unoccupied states, and, therefore, require pseudopotentials that are capable of accurately reproducing scattering states several Rydberg above the Fermi level. Norm-conserving pseudopotentials have been shown to have good excited-state properties,[50] and the addition of multiple projectors per angular momentum channel can improve them further.[51]

```
dft.program qe
calc.mode xas
calc.edges { -22 2 1 }
structure.rprim { 7.379568958 0 0
                  0 7.379568958 0
                  0 0 7.379568958 }
structure.znucl { 38 22 8 }
structure.typat{ 1 2 3 3 3 }
structure.xred {
 0.5  0.5  0.5
 0.0  0.0  0.0
 0.5  0.0  0.0
 0.0  0.5  0.0
 0.0  0.0  0.5 }
```

Fig. 3 A minimal input for calculating the Ti $L_{2,3}$ XAS of $SrTiO_3$ using QUANTUM ESPRESSO as the external DFT program.

Support for optimized norm-conserving Vanderbilt pseudopotentials has been added by extending the oncvpsp code.[26,27] This makes files from two large pseudopotential databases usable, SG15[52] and PseudoDojo.[53] The input files from version 0.4 of PseudoDojo are included in the OCEAN distribution. A database of all these pseudopotentials can be created and installed with OCEAN, allowing to automatically load the correct pseudopotentials based on the elements in the input file and to set the plane-wave cut-off based on the included pseudopotentials. Previously, OCEAN relied on an internal atomic code, and pseudopotentials had to be supplied following the Fritz-Haber-Institut (FHI) format. Some pseudopotentials following this format could be found on the ABINIT or QUANTUM ESPRESSO websites, or they could be generated by the user using the opium code.

3.1.2 Static dielectric constant. For both valence-level and core-level calculations OCEAN makes use of a model dielectric function to calculate part or all of the screening of the direct interaction (see Section 3.2 and ref. 29). The model takes as an input the electronic contribution to the isotropic static dielectric constant ε_∞. In previous versions of OCEAN this was a required input, and the user was encouraged to find a value from experimental measurements or self-consistently with valence OCEAN calculations. As of version 3, OCEAN can calculate the dielectric constant within density–functional perturbation theory using QUANTUM ESPRESSO.[54] For metallic systems, a default of $\varepsilon_\infty = 10\,000$ is used. Accuracy is not needed for conducting systems as the errors scale with the inverse of the dielectric constant. The effect on calculated XAS of incorrect values for the dielectric constant has been discussed elsewhere.[29]

3.1.3 Core-level broadening. A constant broadening term must be added to all calculated spectra to avoid divergences in either the Haydock or GMRES methods. For convenience in X-ray spectra this is set to the broadening from the finite core–hole lifetime. By default OCEAN will use the tabulated recommended widths from Campbell and Papp.[40]

3.1.4 Haydock convergence. The number of Haydock iterations used to generate each calculated spectra can be set at a fixed number or determined automatically. The convergence is determined by comparing the area between two curves, the spectra $\varepsilon_2(\omega)$ generated with n iterations and that generated with $n + m$. If the area between the curves, normalized by the average area under both curves, falls below a threshold (default of 0.001) the calculation is stopped. By default, the spacing between the compared spectra is $m = 5$, a variation of a method used elsewhere.[43]

3.1.5 X-ray photon files. The transition operators for core-level spectra are specified using auxiliary photon files which specify the type of operator (dipole, quadrupole [quad], or NRIXS [qRaman]). Additionally, the files define the photon polarization (neglected for NRIXS) and momentum direction (neglected for dipole). While both the quadrupole and NRIXS calculations depend on the magnitude of the momentum, for quadrupole it is proportional to the photon energy and for NRIXS it is set directly. If no photon files are supplied, OCEAN will automatically generate a set sufficient for calculating the isotropically averaged absorption spectrum using either a dipole or quadrupole operator. The division between dipole and quadrupole operators is made somewhat arbitrarily at 4000 eV, starting at the Ca K edge or the Sn L edge. Energies are taken from ref. 55.

3.2 Screening

Details of the improved method, implementation, and robustness of the screening calculation in OCEAN have been published previously,[29] and the main points will be summarized here. The direct interaction is screened using the random phase approximation (RPA) dielectric response. In OCEAN this calculation is carried out in real space using a combination of the RPA response and a model dielectric function. The RPA response is calculated for a neutral excitation consisting of a core hole or point charge surrounded by a neutralizing shell at some radius r_S. This approximation is controlled by the shell radius with the contribution of the model going to zero as $r_S \to \infty$. This approach can be used for both valence and X-ray excitations.

The screening calculation itself scales well, $N^2 \log N$ where N is system size. However, the calculation requires electron orbitals from a DFT calculation which scales as N^3, and for large calculations the total time required is dominated by the DFT calculation. The screening calculation is parallelized both over atomic sites as well as within a single site.

3.3 Exciton plotting

At times it is useful to plot out the density of the excited electron or hole that make up the exciton to try and gain insight from the shape of the excitation. Unlike codes that diagonalize the BSE Hamiltonian and can therefore plot excitons from the eigenvectors,[11,13,57] OCEAN only calculates an exciton vector which includes a projection based on the photon and a finite energy bandwidth (see eqn (24)). In many cases, however, this exciton vector is the physically relevant quantity to plot. The finite energy smearing from $i\eta$ is set to match the core–hole lifetime broadening and captures the superposition of eigenstates that are populated. The real-space exciton ϕ for a specific $x(\omega)$ can be generated

$$\phi(\mathbf{r}+\mathbf{R};\omega) = \sum_{c\mathbf{k}} \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{r}+\mathbf{R})} u_{c\mathbf{k}}(\mathbf{r}) x_{\gamma c\mathbf{k}}(\omega) \tag{31}$$

where $\mathbf{r}$ is within a unit cell and $\mathbf{R}$ are integer factors of the lattice vectors. The vector $x(\omega)$ is the output from a GMRES calculation.

The excitons are localized and need not be generated on an R-grid the same size as the k-point grid. Additionally, the real-space r-mesh need not be the same as the x-mesh used in the BSE calculation to generate $x(\omega)$. A user can specify an integer multiple supercell or an explicit right rectangular prism. In the case of the latter, the core site is centered and interpolation is used to transform from the lattice-vector aligned $\phi(\mathbf{r},\mathbf{R})$ to a regular orthogonal real-space grid. The exciton density is output in the cube file format, allowing plotting with various third-party tools. As an example, the lowest energy exciton at the nitrogen K edge of a single sheet of hexagonal boron nitride is shown in Fig. 4.

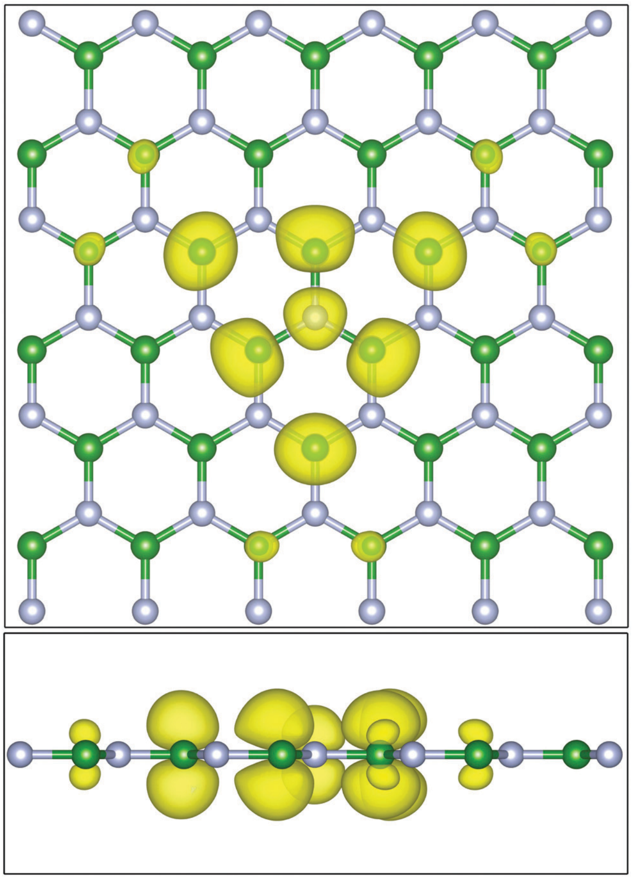

Fig. 4 (Top) The excited electron density (yellow) of the lowest energy exciton of a single sheet of hexagonal boron nitride at the nitrogen K edge viewed looking down. (Bottom) the same plot but viewed from the side. The absorbing site is the central nitrogen atom (grey). The majority of the electron density is sitting in p_z type states on the first and second nearest neighboring boron atoms (green). Plots generated using VESTA.[56]

3.4 Temperature-dependent XAS

At temperatures near or below room temperature, the main effect of temperature on calculated X-ray spectra comes through the motion of the ions. The electronic state occupation numbers can be taken in the $T \to 0$ K limit as either 0 or 1. (As an aid to convergence of the self-consistent DFT calculation, a smearing of the occupation numbers is used to generate the electron density. While this smearing can be treated as a temperature, the goal is to approximate the zero-temperature density while replacing the integral over the Brillouin zone with a coarse sum over k-points.) However, at elevated temperatures or in response to an optical pump, non-integer occupation numbers $0 \leq n \leq 1$ are possible. If the ions are held fixed, modeling either temperatures below the melting point or short-delay pump–probe experiments which measure before the lattice thermalizes, then the two initial effects on the spectra are due to changes in Fermi blocking and changes to the dielectric screening. The primary effect of the latter is an increase in the long-range dielectric screening of insulating or semiconducting systems due to an increase in charge carriers. Only the first of these effects has been implemented in OCEAN and is described below.

Following previous work,[59,60] finite temperature has been approximated by modifying the BSE Hamiltonian to include fractional occupation numbers

$$H_{ij} = \varepsilon_i\delta_{ij} + \sqrt{\tilde{f}_i}[V_X - W]\sqrt{\tilde{f}_j} \qquad (32)$$

where $\tilde{f}_i = |f_{ei} - f_{hi}|$ is the Fermi factor difference between the nominal electron and hole states. The individual Fermi factors are given by the expected $f = [\exp[(\varepsilon - E_F)/k_BT] + 1]^{-1}$, where ε is the energy of the electron or hole state, E_F is the Fermi level, k_B is Boltzmann's constant, and T the temperature. At $T \to 0$, these factors reduce to 1 or 0, and Eq. 32 is equivalent to algorithmically limiting the space of electrons and holes and neglecting the $\sqrt{\tilde{f}}$ factors. Importantly, the Fermi level itself will shift with temperature as the occupied and unoccupied densities of states are unlikely to be symmetric.

To showcase the effect of temperature of XAS, calculations of bulk copper were carried out at both the L and K edges. Previously, a pump-probe experiment was carried out measuring the changes in the Cu L edge in response to a 400 nm laser pump for probe delay times of 0 ps to 21 ps.[58] The laser fluence was sufficient to damage the sample, and previous simulations of the L-edge absorption accounted for this using molecular dynamics to model high-temperature liquid copper.[58,61] There is a short delay between the thermalization of the electronic system and that energy being transferred to phonons, heating and then melting the copper lattice. Here, only the electronic temperature is considered. In Fig. 5, calculations at 300 K and 10 200 K are compared with measurements at room temperature and with a 2 ps delay between the laser pump and X-ray probe. The calculations were broadened by 0.305 eV to account for the L_3 lifetime broadening and an additional 2.5 eV full-width at half maximum (FWHM) Gaussian. The calculations were both shifted by the same amount to fit the onset of the L_3 edge of the room temperature measurement. The redshift with increasing temperature is a direct result of changes to the occupation numbers of states near the Fermi level. The OCEAN calculations agree well with the measured data. The temperature activated features below the L_2 (948 eV) and L_3 (930 eV) are stronger in the measurement. These are due to transitions into the 3d orbitals which lie just below the Fermi level and are occupied at low temperature. The position and distribution of these orbitals from the DFT calculations may not be correct, and changes in their energy will have a large effect on this part of the spectrum.[58] Here, a Hubbard-U correction of 2.0 eV was applied to the copper d orbitals.[48] Realistic simulations of high-temperature XAS can be used to measure the electronic temperature of samples in these types of high-fluence pump probe experiments, but proper accounting of the ionic temperature is also necessary.[58,61–64]

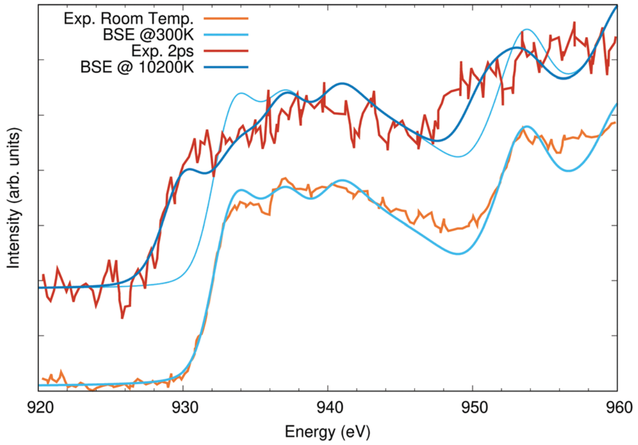

Fig. 5 Cu $L_{2,3}$ edge X-ray absorption calculations compared to measured data taken from ref. 58. The lower spectra were taken and calculated at room temperature, while the upper spectra, offset vertically for clarity, are at elevated temperature (see text for details). The 300 K calculation is shown a second time with the high-temperature spectra to highlight the shifts in spectral weight with temperature.

Finite-temperature effects have also been observed in copper K-edge XAS. A two-photon measurement was carried out using an X-ray laser tuned to half the K-edge energy (4500 eV).[65] Unlike standard absorption measurements, which for 3d transition metal K edges are almost entirely within the dipole limit, this measurement explicitly probed quadrupole-like transitions. The first photon excites an electron from the 1s to a highly-non-resonant virtual p-like orbital, and the second from that p orbital into an energetically-allowed s or d state. The two-photon absorption is calculated by assuming that the two photons are absorbed in quick succession with no relaxation in the short-lived intermediate state.

The transition matrix element for this process, following eqn (20), is given by

$$M_{c\mathbf{k};\varepsilon} \propto \sum_{\nu lm} \langle c\mathbf{k}|\nu lm\rangle\langle \nu lm|\hat{\varepsilon}\cdot\mathbf{r} \sum_{n\geq 4} \frac{|np\rangle\langle np|}{\omega - (\varepsilon_{np} - \varepsilon_{1s})}\hat{\varepsilon}\cdot\mathbf{r}|1s\rangle \qquad (33)$$

where the absorption is taken to be dipole limited and the 2p and 3p orbitals of copper are filled. The energy of the X-ray photon ω = 4500 eV is approximately half the difference between the energy 1s core level ε_{1s} and the low-lying unoccupied p orbitals ε_{np}. While the sum over n is infinite, the matrix elements with the 1s level fall off. If the sum is limited to excitations within even a few 100 eV of the Fermi level, the 4500 eV detuning of the X-ray makes the energy denominator nearly independent of n. It is therefore approximated as a constant. Following that approximation, the transition into any unoccupied p-like state can be replaced with one minus the occupied 2p and 3p core levels, giving

$$M_{c\mathbf{k};\varepsilon} \propto \sum_{\nu lm} \langle c\mathbf{k}|\nu lm\rangle\langle \nu lm|\hat{\varepsilon}\cdot\mathbf{r}[1 - |2\mathrm{p}\rangle\langle 2\mathrm{p}| - |3\mathrm{p}\rangle\langle 3\mathrm{p}|]\hat{\varepsilon}\cdot\mathbf{r}|1\mathrm{s}\rangle \qquad (34)$$

In contrast to a quadrupole transition there is an intermediate excited state and a second $\hat{\varepsilon}\cdot\mathbf{r}$ takes the place of the photon momentum $\mathbf{q}\cdot\mathbf{r}$. For a laser both photons will have identical polarizations, $\hat{\varepsilon}\|\hat{\varepsilon}$, while for photons the momentum must be perpendicular, $\hat{\varepsilon}\perp\mathbf{q}$. Like the standard X-ray absorption matrix elements, the terms in eqn (34) were calculated using real-space integration of the all-electron OPFs and core-level orbitals from the atomic all-electron calculation.

In Fig. 6, I show the measured data reproduced from ref. 65 for both standard XAS and two-photon XAS. The two-photon XAS measurement was not explicitly at elevated temperature. However, the high X-ray fluence required for the measurement resulted in inevitable sample heating, and the large 6 eV redshift is evidence of a significant depopulation of the 3d bands. The measured two-photon absorption is not consistent with a single elevated electronic temperature. The redshift can be reproduced with temperatures exceeding 42 000 K but at this temperature a significant white line is predicted. This extreme temperature is also inconsistent with the experimental setup or measurements of the one-photon absorption which show only a slight decrease in intensity around 8980 eV,[65] reflecting only a small amount of Fermi blocking due to thermal excitations. Instead, the observed two-photon absorption is consistent with non-thermal depopulation of the 3d band, likely either due to one-photon absorption from the 3d orbitals or far-from resonance excitations from the $L_{2,3}$ followed by radiative or non-radiative decays involving the 3d states. To approximate this, a constant occupation of 0.91 has been assumed for valence bands (4s and 3d) below the Fermi level. In Fig. 6, calculations using either this reduced valence occupation or 300 K Fermi statistics are compared to the measured two-photon and one-photon absorption, respectively, including an additional 1.5 eV FWHM of Gaussian broadening. The calculated two-photon absorption shows decent agreement with the measured data, but it may be fortuitous. The data is sparse and has relatively large uncertainty in the intensity.

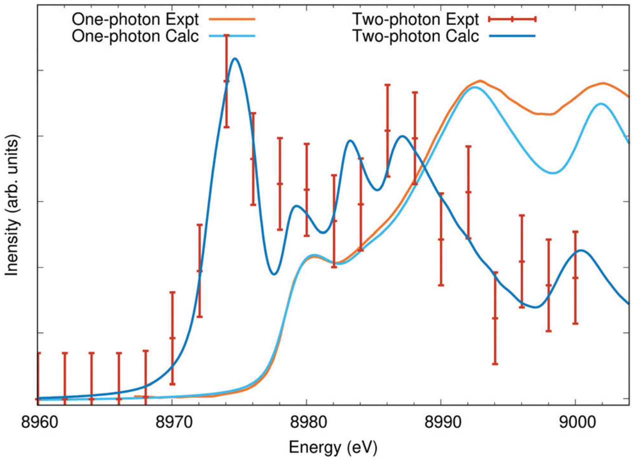

Fig. 6 Cu K edge X-ray absorption calculations compared to measured data taken from ref. 65 (error bars from the original publication). The two-photon calculated spectra assume a 9% depopulation of the valence bands, leading to the appearance of a strong pre-edge peak from transitions into the otherwise fully occupied 3d states.

4 Further examples

4.1 Dependence on the underlying DFT

To highlight the effect of the exchange–correlation potential on computed spectra, the O and Fe edges of α-Fe_2O_3 hematite have been calculated using three different approximations: LDA, PBE, and SCAN. The lattice parameters for α-Fe_2O_3 hematite were taken from experiment,[67] but for each functional the atoms were allowed to relax. No Hubbard-U parameter was included in the calculations. While a $+U$ parameter is capable of reproducing the band gap, it also has the effect of collapsing the two distinct peaks in the O K-edge XAS into a single feature. The band gaps were calculated to be approximately 0.40 eV with LDA, 0.69 eV with PBE, and 1.55 eV with SCAN, as compared to an experimental value of approximately 2 eV.[66] The calculated magnetic moments on the Fe sites were found to be $3.4\mu_B$, $3.4\mu_B$, and $3.7\mu_B$, respectively.

The Fe $L_{2,3}$ edge is shown in Fig. 7(a), and all three calculations are normalized and aligned to match the first peak at 707.5 eV. Among the three calculated spectra, only very slight differences in the crystal field splitting is evident between the 3d orbitals: the t_{2g} at 707.5 eV and the e_g 709 eV and again at the L_2 edge at 721 eV and 722.5 eV. However, there is a significant difference in the spectral weight distribution between the t_{2g} and e_g, with PBE giving the smallest e_g and SCAN the largest. It is not clear which gives the best agreement with experiment. The BSE does not capture the full multiplet structure of the Fe $L_{2,3}$ edge and is missing coupling to secondary dd* excitations.[68,69] The calculated spin–orbit splitting between the Fe $2p_{3/2}$ and $2p_{1/2}$ states was scaled by 1.09 to better match the observed spacing between the L_3 and L_2 edges. The O K edge in Fig. 7(b) provides another view of the same orbitals. The first two peaks in the oxygen XAS of transition metal oxides are due to hybridization between O p and Fe d orbitals. Here the ratio between t_{2g} and e_g clearly favors the LDA and PBE over the SCAN calculation, an outcome counter to the performance of the three functionals in reproducing the band gap.

4.2 State-dependent lifetime broadening

In a number of systems, the lifetime broadening of the excited state cannot be taken as a constant across the spectra. In general, the photoelectron lifetime gets shorter with higher excitation energy, and the broadening steadily increases over an energy scale set by the plasmon energy. It has also been found that valence-band lifetimes can vary widely with orbital, and both conduction and valence band lifetime effects can be included in OCEAN calculations through a complex-valued self-energy correction.[18] At $L_{2,3}$ edges, the $2p_{3/2}$ holes have longer lifetimes than the $2p_{1/2}$ holes due to Coster-Kronig transitions from a $2p_{1/2}$ hole to a $2p_{3/2}$ but not *vice versa*. As noted in Section 2.2.1.3, this effect can be captured by including an imaginary component to the spin–orbit splitting.

To showcase this, I examine calculations of the titanium $L_{2,3}$ edge of strontium titanate and compare to experiment in Fig. 8.

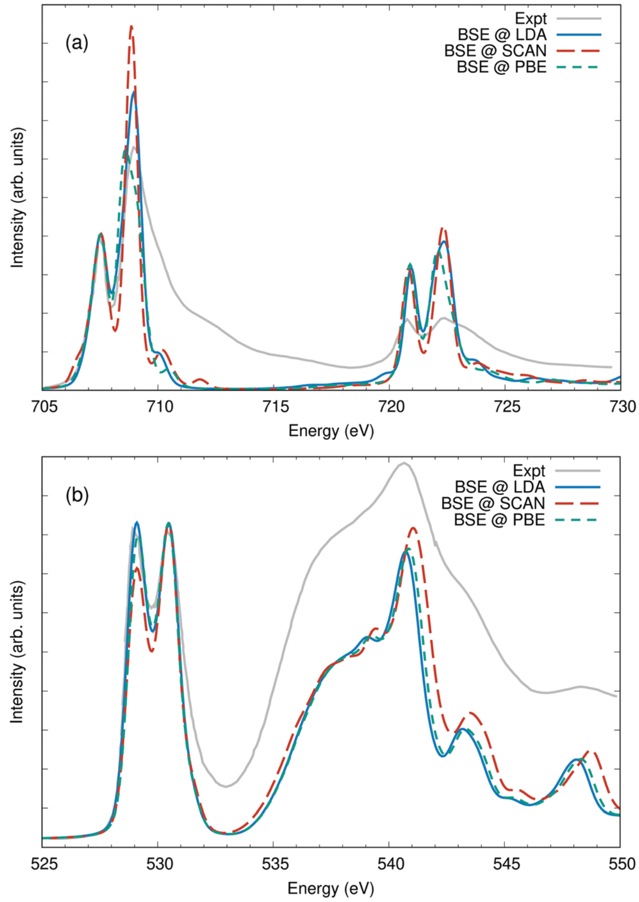

Fig. 7 (a) The iron $L_{2,3}$ edge and (b) the oxygen K edge of Fe_2O_3 calculated using three different approximations to the DFT exchange correlation potential and compared with experiment from ref. 66. The iron L edges were aligned and scaled to the lowest feature at 707.5 eV and the O K edge to the second fearture at 530.5 eV. The differences between the three calculations are minor, however the SCAN potential shows a markedly higher intensity in the e_g orbitals compared to PBE or LDA as can be noted at 530.5 eV or 709 eV.

The spectra splits into four main peaks. The empty 3d orbitals are divided by symmetry into two groups, the lower-energy t_{2g} (d_{xy}, d_{xz}, d_{yz}) which minimize overlap with the neighboring oxygen atoms and the higher-energy e_g ($d_{x^2-y^2}$, d_{z^2}) which point along the Ti–O bonds. These two 2p → 3d transitions are split into four by the 2p spin–orbit splitting. The Ti 2p spin–orbit splitting parameter ζ_{2p} is calculated to be 3.83 eV, but was lowered to 3.76 eV to better match experiment. The L_3 and L_2 core–hole lifetime broadening were set to 0.25 eV and 0.52 eV, respectively.[40] An additional 0.5 eV broadening factor was added to the the e_g orbitals to match the additional vibrational broadening of these states.[71] The calculated spectra show agreement with the measured data in line with previous first-principles calculations.[70,72,73]

4.3 Finite momentum transfer

As mentioned in Section 2.3.1.2, OCEAN can calculate non-resonant inelastic X-ray scattering (NRIXS) in which an X-ray photon scatters off of a core-level electron transferring both energy and momentum to the electron. Experimentally, NRIXS gives additional flexibility to X-ray measurements. Soft X-ray transitions below 1000 eV can be probed using hard X-rays with much longer penetration depths. The longer penetration depth allows a variety of sample conditions not possible with soft X-ray such as measuring at high pressures or measuring a component of an operating device like a battery or catalyst. NRIXS also decouples the energy and momentum transfer of the excitation. Measurements at a single X-ray edge can be carried out for several momenta, revealing dipole-forbidden transitions and allowing an investigation of the momentum dependence of the local density of states around the absorbing atom. Here, OCEAN calculations of liquid water have been carried out both in the dipole limit and for several finite momenta.

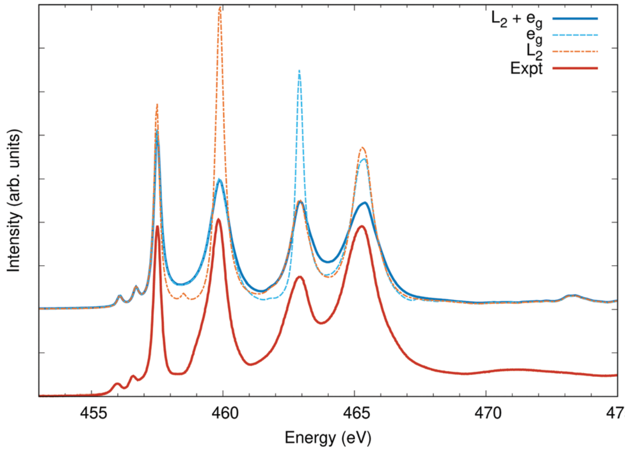

Fig. 8 The Ti $L_{2,3}$ edge of $SrTiO_3$ calculated using three different broadening methods compared to measured data taken from ref. 70. All calculated spectra have a uniform broadening of 0.25 eV full-width at half maximum. Selectively, peaks associated with the L_2 edge ($2p_{1/2}$ holes) are broadened to 0.52 eV, and peaks associated with the 3d e_g states are broadened by an additional 0.5 eV.

The near-edge X-ray spectra of both liquid water and ice have been studied extensively over the past decade. It is known that the hydrogen bonding network in water affords both low density and high density phases (in part due to the existence of both low-density and high-density amorphous ice). However, debate continues around the coexistence of low and high density liquid water phases and the nature of density fluctuations in liquid water.[75–78] Large simulation cells are necessary to capture proposed density fluctuations that reach the scale of ≈1 nm.[76] First-principles calculations of the oxygen K-edge spectra have been carried out, including using the BSE,[17,31,74,79] but often on small unit cells that might have large artifacts due to artificial periodicity or using various additional approximations to the BSE interactions.

Here, 64-molecule water cells were generated by the Deep Potential Molecular Dynamics model.[81–83] This model was trained using the PBE0 functional[84] and Tkatchenko-Scheffler approximation to the van der Waals interactions.[85] The snapshots were generated using path-integral molecular dynamics with 8 beads and NPT conditions: 64 molecules, 300 K, and 100 kPa. (1 bar).[86] OCEAN calculations were carried out using the

SCAN functional and the LDA approximation to f_{xc} for the core–hole screening[29] on 16 snapshots: two time steps separated by 1 ps across all 8 beads. For each snapshot spectra are calculated for 3 orthogonal photon polarizations (for XAS) or momenta (NRIXS) and averaged together to form a single σ_i. These are averaged to form a single average spectrum $\bar{\sigma}$ with a variance S given by $S^2(\omega) = \sum_i^N (\bar{\sigma}(\omega) - \sigma_i(\omega))^2/N$. The resulting spectrum, broadened by an 0.6 eV FWHM Gaussian in addition to the 0.08 eV of core–hole broadening, is shown in Fig. 9 compared with experiment taken from ref. 74.

The photon-momentum dependence is shown in Fig. 10. Again the calculations were aligned to the onset of the main edge, resulting in an 0.3 eV relative shift with respect to Fig. 9. Here the intensity is normalized to the main edge region, between 537 eV and 543 eV, and compared to measurements at q = 2.6 Å^{-1} and q = 8.2 Å^{-1}.[75] The agreement at different momenta and in comparison to different experimental data sets is comparable. Furthermore, by contrasting low-q and high-q NRIXS, it is clear that the relative balance between s-type and p-type final states around the oxygen atoms is similar between the calculation and experiment, see Fig. 10(b). The same discrepancies are seen between the calculation and measurements at all three values of the momentum transfer: a weaker pre-edge at 535 eV, slight overstatement of the main edge at 537 eV, and too narrow overall resulting in reduced spectral weight around 544 eV. It is possible GW corrections would alleviate some of this, though recent GW-BSE calculations show a comparable level of agreement, albeit using a different DFT exchange–correlation and different approximations to the electron–hole interactions.[79] It would be instructive to compare calculations of various ice phases to understand if these deficiencies arise due to the structural models used for liquid water or if they are more general to water and ice systems.

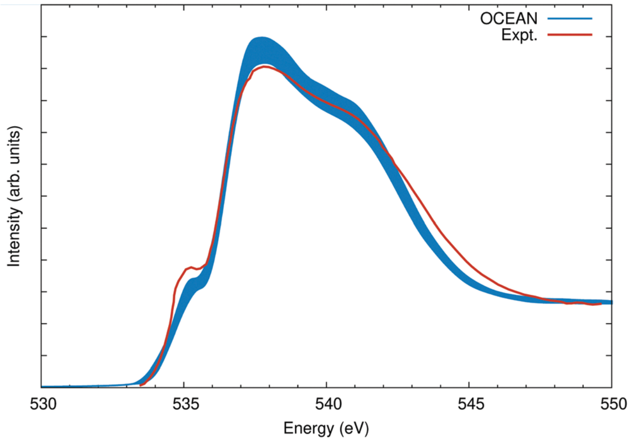

Fig. 9 The O K edge of liquid water calculated as an average over 16 snapshots and compared to measurements from ref. 74. The width of the OCEAN line reflects the variance in the mean (see text). Both the calculation and experiment are non-resonant X-ray scattering at a momentum transfer of 3.1 Å^{-1}, which is nearly the same as the dipole absorption [see Fig. 10(b)].

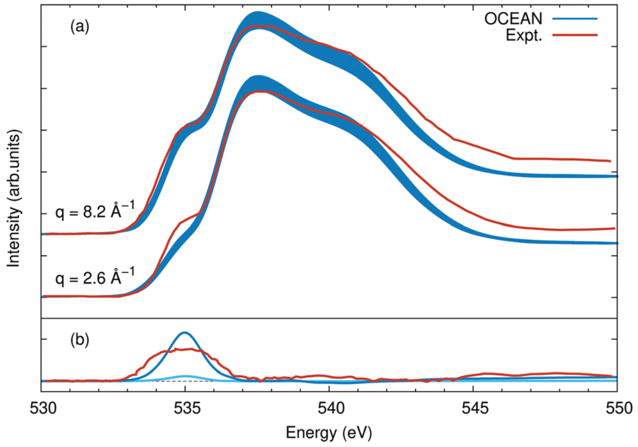

Fig. 10 (a) The O K edge of liquid water as seen with NRIXS at momenta of 2.6 Å^{-1} and 8.2 Å^{-1}. Other than the value of the momentum transfer the details of the calculation are the same as for Fig. 9, and the experimental data is from ref. 75. (b) The difference between the high and low momentum spectra for OCEAN in blue and the measured data in red. For reference, the difference between q = 2.6 Å^{-1} and a dipole XAS calculation is shown in cyan, showing already a slight increase in the pre-edge feature at finite momentum. While the calculated pre-edge appears to be growing faster with increasing momentum in (a) than the measured pre-edge, in (b) it is clear that in the measured data the pre-edge is wider at higher q. The change in the area of the pre-edge is very similar between OCEAN and experiment.

4.4 Timing

A practical concern with any computational method is the cost – the product of the amount of time and size of computer system required to carry out a calculation. BSE methods have long been considered expensive. The fundamental scaling of the BSE is N^3 where N is a measure of the unit cell size. However, the Haydock recursion is only the final step in a BSE calculation, and, despite its poor scaling with system size, usually not the longest step. Pre-computing the dielectric screening for the direct interaction scales as $N^2 \log N$,[29] and the DFT calculations to generate orbitals for both the screening and BSE calculations scale as N^3.

The time required to run several of the examples shown so far using a desktop computer[3,80] is summarized in Table 1. These are all small systems, under 700 a.u.^{3} and with 1 to 10 atoms in the unit cell. To compare with previous performance, the timing using OCEAN version 2.0.3 are also shown for hBN and SrTiO$_3$. The DFT times between code versions are not directly comparable as the older version of OCEAN requires an older version of QUANTUM ESPRESSO (5.2.0) and does not support the same PseudoDojo pseudopotentials. However, the times for the DFT stage are similar. The largest difference in timing is seen in the Screen stage, but the time required for the BSE stage has also been substantially reduced.

To test the performance of OCEAN for large systems the oxygen K edge of water was calculated using 64-, 128-, and 256-molecule cells using a small computer cluster.[3,87] The cells were constructed by making super cells of the aforementioned 64-molecule water snapshots. The computational cost of

Table 1 The time required for running a selection of X-ray absorption calculations on a single workstation.[80] Timings for the previous version of OCEAN are provided for h-BN and $SrTiO_3$

System	Edge	Vol. (a.u.^{3})	N_k	N_b BSE	N_b screen	DFT (s)	Prep (s)	Screen (s)	BSE (s)	Total (min)
h-BN	N K	390.0	$16 \times 16 \times 4$	36	271	523.6	42.0	163.2	68.1	13.3
* v2.0.3						628.7	53.1	913.4	120.8	28.6
$SrTiO_3$	Ti $L_{2,3}$	401.9	$8 \times 8 \times 8$	128	300	777.9	34.8	350.3	115.2	21.3
* v2.0.3						760.0	49.3	1827.3	161.1	46.6
Cu	Cu K	79.4	$16 \times 16 \times 16$	10	90	146.8	52.5	277.2	36.1	8.5
Fe_2O_3	Fe $L_{2,3}$	678.8	$8 \times 8 \times 8$	60	257	9055.8	196.3	72.0	85.5	156.8
CdS	S $L_{2,3}$	674.3	$12 \times 12 \times 8$	40	256	4576.9	136.8	415.3	247.6	89.6

Table 2 The time required for calculations of the O K edge of liquid water using various sized cells on a small computer cluster.[87] The values given for the 64-molecule cell are averages over several runs (see text). For the 256-molecule cell, thread parallelization was also used for the DFT to reduce the total memory usage

N_{H_2O}	N_{proc}	DFT (m)	Prep (m)	Screen (m)	BSE (m)
64	256	36.1	4.7	7.6	20.2
128	256	173.2	7.9	24.8	60.4
256	512	576.3	23.2	60.6	162.9

calculating the spectra of various sized water cells is summarized in Table 2. These calculations show that OCEAN is suitable for simulating cells up to 50 000 a.u.^{3} or around 7 nm^3.

5 Outlook and future development

The OCEAN package provides an easy-to-use interface for calculating spectra using the Bethe-Salpeter equation approach. Recent advances have focused on improving the reliability and usefulness of the code for end users by simplifying the inputs and reducing the required information to a minimum. Here a few examples have been shown highlighting new functionality that has been added to OCEAN. Future development of OCEAN will focus in part on increased usability, such as additional pre- and post-processing scripts to facilitate converting structural information such as the Crystallographic Information File (CIF) to an OCEAN input as well as automatically calculating additional ground-state properties such as the projected density of states to aid in interpreting spectra. The DFT stage of OCEAN takes a substantial part of the total computational time due to the large number of unoccupied states required for both the BSE and the screening calculations. Various methods for extrapolation, wave function approximation, or exploiting completeness relations have been suggested in literature and should be investigated for incorporation into OCEAN.

There are a number of approximations in OCEAN that can be alleviated through future development. The Tamm-Dancoff approximation is not mandatory in several valence-level BSE codes,[11,12] and going beyond it may prove important for simulating momentum-dependent RIXS. The valence BSE solver is now capable of calculating interactions of triplet states in support of $L_{2,3}$ RIXS, but OCEAN does not support directly calculating spin-triplet valence-level exciton states. For heavier elements or surface states the valence and conduction band spin–orbit coupling can become significant, but currently OCEAN requires collinear spin in the DFT orbitals. Finally, moving beyond the static screening approximation is required to more realistically simulate finite-temperature or non-equilibrium ground states (pump–probe experiments), especially in materials that are semiconductors or insulators in their ground state. Dynamic screening may also be important for transition metal L edges where the 2p splitting is of the same order as the plasmon response, but frequency dependent screening has not yet been investigated for core-level BSE spectra.

At present, OCEAN is limited to simulating excitations that consist of a single electron–hole pair. This ignores the coupling to secondary electron–hole excitations, such as what is seen in the multiplet structure of transition metal $L_{2,3}$ edges with partial d-band occupancy.[39] It also precludes phonon dynamics in response to the creation of the core hole which can influence both absorption and emission spectra. The challenge is extending the BSE approach without a large increase in the size of the Hamiltonian such that calculating spectra of extended systems remains tractable. Some work has been done using a cumulant spectral function to add many-body effects to BSE spectra.[70] However, this approach simplifies the spectral function as independent of the photoelectron, and it is not currently applicable if the many-body effects are dependent on symmetry or localization of the photoelectron.

Author contributions

J. V. wrote the manuscript and implemented the changes to the OCEAN code.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The development of what is now OCEAN has taken place over more than 20 years and long predates my involvement and the name OCEAN itself. Much of the work has been carried out by Eric L. Shirley, but I would also like to acknowledge the various contributors and coauthors of earlier versions and earlier efforts: L. X. Benedict, R. B. Bohn, S. D. Dalosto, K. Gilmore, J. J. Kas, A. Krotz, H. M. Lawler, Z. H. Levine, Y. Liang, S. D. Pemmaraju, D. Prendergast, J. J. Rehr, J. A. Soininen, and F. Vila.

Notes and references

1 J. Vinson, J. J. Rehr, J. J. Kas and E. L. Shirley, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, **83**, 115106.

2 K. Gilmore, J. Vinson, E. Shirley, D. Prendergast, C. Pemmaraju, J. Kas, F. Vila and J. Rehr, *Comput. Phys. Comm.*, 2015, **197**, 109–117.

3 Specific software and hardware are identified for information purposes. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the software or hardware identified are necessarily the best available for the purpose.

4 H. M. Lawler, J. J. Rehr, F. Vila, S. D. Dalosto, E. L. Shirley and Z. H. Levine, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **78**, 205108.

5 The ocean code is available at **https://www.ocean-code.com** v. 3.0.0.

6 M. Rohlfing and S. G. Louie, *Phys. Rev. Lett.*, 1998, **80**, 3320–3323.

7 L. X. Benedict, E. L. Shirley and R. B. Bohn, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **57**, R9385–R9387.

8 L. X. Benedict, E. L. Shirley and R. B. Bohn, *Phys. Rev. Lett.*, 1998, **80**, 4514–4517.

9 E. L. Shirley, *Phys. Rev. Lett.*, 1998, **80**, 794–797.

10 X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J.-M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J.-B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A.-L. Marques, A. Martin, C. Martins, H. P.-C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G.-M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. V. Troeye, M. J. Verstraete, G. Zérah and J. W. Zwanziger, *Comput. Phys. Commun.*, 2020, **248**, 107042.

11 D. Sangalli, A. Ferretti, H. Miranda, C. Attaccalite, I. Marri, E. Cannuccia, P. Melo, M. Marsili, F. Paleari, A. Marrazzo, G. Prandini, P. Bonfà, M. O. Atambo, F. Affinito, M. Palummo, A. Molina-Sánchez, C. Hogan, M. Grüning, D. Varsano and A. Marini, *J. Phys.: Condens. Matter*, 2019, **31**, 325902.

12 J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen and S. G. Louie, *Comput. Phys. Commun.*, 2012, **183**, 1269–1289.

13 C. Vorwerk, B. Aurich, C. Cocchi and C. Draxl, *Electronic Structure*, 2019, **1**, 037001.

14 F. M. de Groot, H. Elnaggar, F. Frati, R. Pan Wang, M. U. Delgado-Jaime, M. van Veenendaal, J. Fernandez-Rodriguez, M. W. Haverkort, R. J. Green, G. van der Laan, Y. Kvashnin, A. Hariki, H. Ikeno, H. Ramanantoanina, C. Daul, B. Delley, M. Odelius, M. Lundberg, O. Kuhn, S. I. Bokarev, E. Shirley, J. Vinson, K. Gilmore, M. Stener, G. Fronzoni, P. Decleva, P. Kruger, M. Retegan, Y. Joly, C. Vorwerk, C. Draxl, J. Rehr and A. Tanaka, *J. Electron Spectrosc. Relat. Phenom.*, 2021, **249**, 147061.

15 in *International Tables for Crystallography Volume I, X-ray absorption spectroscopy and related techniques*, ed. C. T. Chantler, F. Boscherini and B. Bunker, 2020.

16 C. Brouder, D. Cabaret, A. Juhin and P. Sainctavit, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **81**, 115125.

17 H. Guo, E. Strelcov, A. Yulaev, J. Wang, N. Appathurai, S. Urquhart, J. Vinson, S. Sahu, M. Zwolak and A. Kolmakov, *Nano Lett.*, 2017, **17**, 1034–1041.

18 J. Vinson, T. Jach, M. Müller, R. Unterumsberger and B. Beckhoff, *Phys. Rev. B*, 2016, **94**, 035163.

19 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.

20 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.

21 P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. D. Corso, S. de Gironcoli, P. Delugas, R. A.-D. Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. O. de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu and S. Baroni, *J. Phys.: Condens. Matter*, 2017, **29**, 465901.

22 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

23 J. Sun, A. Ruzsinszky and J. P. Perdew, *Phys. Rev. Lett.*, 2015, **115**, 036402.

24 L. Hedin, *Phys. Rev.*, 1965, **139**, A796–A823.

25 J. Vinson and J. J. Rehr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2012, **86**, 195135.

26 D. R. Hamann, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **88**, 085117.

27 The open-source code oncvpsp is available at **http://www.mat-simresearch.com** v. 3.3.1. Modifications for use with ocean are avaiable at **https://github.com/jtv3/oncvpsp**.

28 P. E. Blöchl, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**, 17953–17979.

29 J. Vinson and E. L. Shirley, *Phys. Rev. B*, 2021, **103**, 245143.

30 W. Zhang, M. Topsakal, C. Cama, C. J. Pelliccione, H. Zhao, S. Ehrlich, L. Wu, Y. Zhu, A. I. Frenkel, K. J. Takeuchi, E. S. Takeuchi, A. C. Marschilok, D. Lu and F. Wang, *J. Am. Chem. Soc.*, 2017, **139**, 16591–16603.

31 J. Vinson, J. J. Kas, F. D. Vila, J. J. Rehr and E. L. Shirley, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2012, **85**, 045101.

32 M. Wansleben, J. Vinson, A. Wählisch, K. Bzheumikhova, P. Hönicke, B. Beckhoff and Y. Kayser, *J. Anal. At. Spectrom.*, 2020, **35**, 2679–2685.

33 E. E. Salpeter and H. A. Bethe, *Phys. Rev.*, 1951, **84**, 1232–1242.

34 G. Onida, L. Reining and A. Rubio, *Rev. Mod. Phys.*, 2002, **74**, 601–659.

35 X. Blase, I. Duchemin, D. Jacquemin and P.-F. Loos, *J. Phys. Chem. Lett.*, 2020, **11**, 7371–7382.

36 T. Sander, E. Maggio and G. Kresse, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2015, **92**, 045209.

37 M. Rohlfing and S. G. Louie, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **62**, 4927–4944.

38 L. X. Benedict and E. L. Shirley, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**, 5441–5451.

39 F. de Groot, *Coord. Chem. Rev.*, 2005, **249**, 31–63.
40 J. Campbell and T. Papp, *At. Data Nucl. Data Tables*, 2001, **77**, 1–56.
41 R. Haydock, V. Heine and M. J. Kelly, *J. Phys. C: Solid State Phys.*, 1975, **8**, 2591–2605.
42 C. Lanczos, *J. Res. Natl. Bur. Stand.*, 1952, **49**, 33.
43 C. Gougoussis, M. Calandra, A. P. Seitsonen and F. Mauri, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **80**, 075102.
44 Y. Saad and M. H. Schultz, *SIAM*, 1986, **7**, 856–869.
45 T. Aoki and K. Ohno, *Phys. Rev. B*, 2019, **100**, 075149.
46 C. Vorwerk, F. Sottile and C. Draxl, *Phys. Rev. Res.*, 2020, **2**, 042003.
47 L. Weinhardt, O. Fuchs, A. Fleszar, M. Bär, M. Blum, M. Weigand, J. D. Denlinger, W. Yang, W. Hanke, E. Umbach and C. Heske, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2009, **79**, 165305.
48 M. Cococcioni and S. de Gironcoli, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2005, **71**, 035105.
49 L. Weinhardt, O. Fuchs, E. Umbach, C. Heske, A. Fleszar, W. Hanke and J. D. Denlinger, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2007, **75**, 165207.
50 E. Luppi, H.-C. Weissker, S. Bottaro, F. Sottile, V. Veniard, L. Reining and G. Onida, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2008, **78**, 245124.
51 P. E. Blöchl, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1990, **41**, 5414–5416.
52 M. Schlipf and F. Gygi, *Comput. Phys. Commun.*, 2015, **196**, 36–44.
53 M. van Setten, M. Giantomassi, E. Bousquet, M. Verstraete, D. Hamann, X. Gonze and G.-M. Rignanese, *Comput. Phys. Commun.*, 2018, **226**, 39–54.
54 P. Giannozzi, S. de Gironcoli, P. Pavone and S. Baroni, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1991, **43**, 7231–7242.
55 W. Elam, B. Ravel and J. Sieber, *Radiat. Phys. Chem.*, 2002, **63**, 121–128.
56 K. Momma and F. Izumi, *J. Appl. Crystallogr.*, 2011, **44**, 1272–1276.
57 A. Gulans, S. Kontur, C. Meisenbichler, D. Nabok, P. Pavone, S. Rigamonti, S. Sagmeister, U. Werner and C. Draxl, *J. Phys.: Condens. Matter*, 2014, **26**, 363202.
58 B. I. Cho, K. Engelhorn, A. A. Correa, T. Ogitsu, C. P. Weber, H. J. Lee, J. Feng, P. A. Ni, Y. Ping, A. J. Nelson, D. Prendergast, R. W. Lee, R. W. Falcone and P. A. Heimann, *Phys. Rev. Lett.*, 2011, **106**, 167601.
59 A. Schleife, C. Rödl, F. Fuchs, K. Hannewald and F. Bechstedt, *Phys. Rev. Lett.*, 2011, **107**, 236405.
60 D. Sangalli, S. Dal Conte, C. Manzoni, G. Cerullo and A. Marini, *Phys. Rev. B*, 2016, **93**, 195205.
61 T. S. Tan, J. J. Kas and J. J. Rehr, *Phys. Rev. B*, 2021, **104**, 035144.
62 N. Jourdain, L. Lecherbourg, V. Recoules, P. Renaudin and F. Dorchies, *Phys. Rev. B*, 2018, **97**, 075148.
63 B. I. Cho, T. Ogitsu, K. Engelhorn, A. A. Correa, Y. Ping, J. W. Lee, L. J. Bae, D. Prendergast, R. W. Falcone and P. A. Heimann, *Sci. Rep.*, 2016, **6**, 18843.
64 B. Mahieu, N. Jourdain, K. Ta Phuoc, F. Dorchies, J.-P. Goddet, A. Lifschitz, P. Renaudin and L. Lecherbourg, *Nat. Commun.*, 2018, **9**, 3276.
65 K. Tamasaku, E. Shigemasa, Y. Inubushi, I. Inoue, T. Osaka, T. Katayama, M. Yabashi, A. Koide, T. Yokoyama and T. Ishikawa, *Phys. Rev. Lett.*, 2018, **121**, 083901.
66 L. W. Finger and R. M. Hazen, *J. Appl. Phys.*, 1980, **51**, 5362–5367.
67 S. Shen, J. Zhou, C.-L. Dong, Y. Hu, E. N. Tseng, P. Guo, L. Guo and S. S. Mao, *Sci. Rep.*, 2014, **4**, 6627.
68 P. S. Miedema and F. M. de Groot, *J. Electron Spectrosc. Relat. Phenom.*, 2013, **187**, 32–48.
69 P. Kuiper, B. G. Searle, P. Rudolf, L. H. Tjeng and C. T. Chen, *Phys. Rev. Lett.*, 1993, **70**, 1549–1552.
70 J. C. Woicik, C. Weiland, C. Jaye, D. A. Fischer, A. K. Rumaiz, E. L. Shirley, J. J. Kas and J. J. Rehr, *Phys. Rev. B*, 2020, **101**, 245119.
71 F. M.-F. de Groot, J. C. Fuggle, B. T. Thole and G. A. Sawatzky, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1990, **41**, 928–937.
72 P. Krüger, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **81**, 125121.
73 R. Laskowski and P. Blaha, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2010, **82**, 205104.
74 J. Niskanen, C. J. Sahle, K. Gilmore, F. Uhlig, J. Smiatek and A. Föhlisch, *Phys. Rev. E*, 2017, **96**, 013319.
75 A. Nilsson, D. Nordlund, I. Waluyo, N. Huang, H. Ogasawara, S. Kaya, U. Bergmann, L.-A. Näslund, H. Öström, P. Wernet, K. Andersson, T. Schiros and L. Pettersson, *J. Electron Spectrosc. Relat. Phenom.*, 2010, **177**, 99–129.
76 C. Huang, K. T. Wikfeldt, T. Tokushima, D. Nordlund, Y. Harada, U. Bergmann, M. Niebuhr, T. M. Weiss, Y. Horikawa, M. Leetmaa, M. P. Ljungberg, O. Takahashi, A. Lenz, L. Ojamäe, A. P. Lyubartsev, S. Shin, L. G.-M. Pettersson and A. Nilsson, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 15214–15218.
77 J. Niskanen, M. Fondell, C. J. Sahle, S. Eckert, R. M. Jay, K. Gilmore, A. Pietzsch, M. Dantz, X. Lu, D. E. McNally, T. Schmitt, V. Vaz da Cruz, V. Kimberg, F. Gel'mukhanov and A. Föhlisch, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 4058–4063.
78 L. G.-M. Pettersson, Y. Harada and A. Nilsson, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 17156–17157.
79 F. Tang, Z. Li, C. Zhang, S. G. Louie, R. Car, D. Y. Qiu and X. Wu, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2201258119.
80 MacPro (2017) 3 GHz 10-Core Intel Xeon W with 64 GB RAM. All runs carried out using 8 processors.
81 L. Zhang, J. Han, H. Wang, R. Car and W. E, *Phys. Rev. Lett.*, 2018, **120**, 143001.
82 J. Han, L. Zhang, R. Car and W. E, *Commun. Comput. Phys.*, 2018, **23**, 629–639.
83 H.-Y. Ko, L. Zhang, B. Santra, H. Wang, W. E, R. A. DiStasio Jr and R. Car, *Mol. Phys.*, 2019, **117**, 3269–3281.
84 J. P. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys.*, 1996, **105**, 9982–9985.
85 A. Tkatchenko and M. Scheffler, *Phys. Rev. Lett.*, 2009, **102**, 073005.
86 M. Calegari and R. Car, private communication.
87 Each node of the cluster has dual 2.1 GHz 16-Core Xeon Silver 4216 with 96 GB RAM.