Fast, efficient, and accurate dielectric screening using a local real-space approach

John Vinson [1,*] and Eric L. Shirley [2]

[1]*Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*
[2]*Physical Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA*

Various many-body perturbation theory techniques for calculating electron behavior rely on W, the screened Coulomb interaction. Computing W requires complete knowledge of the dielectric response of the electronic system, and the fidelity of the calculated dielectric response limits the reliability of predicted electronic and structural properties. As a simplification, calculations often begin with the random-phase approximation (RPA). However, even RPA calculations are costly and scale poorly, typically as N^4 (N representing the system size). A local approach has been shown to be efficient while maintaining accuracy for screening core-level excitations [Ultramicroscopy **106**, 986 (2006)]. We extend this method to valence-level excitations. We present improvements to the accuracy and execution of this scheme, including reconstruction of the all-electron character of the pseudopotential-based wave functions, improved $N^2 \log N$ scaling, and a parallelized implementation. We discuss applications to Bethe-Salpeter equation calculations of core and valence spectroscopies.

I. INTRODUCTION

For condensed matter systems, one of the fundamental properties is the dielectric response. This is closely related to the polarizability, the movement of the constituent, electrically charged electrons and ions that is responsible for any difference between the applied and total potentials. Here we limit our investigation to the electronic behavior of the system, treating the ions as stationary. The dielectric response determines conductivity as well as frequency-dependent absorption and transmission of photons. The polarizability is used in many-body perturbation theory to more accurately determine electronic properties. In calculations of a variety of properties of condensed matter systems, band alignment, optical absorption, adsorption energies, etc., the determination of the electronic response plays a vital role.

Calculations of electron polarizability in condensed systems are commonplace. For a periodic system (infinite crystal), the polarizability within the random-phase approximation (RPA) is typically calculated in reciprocal space based on the spectral representation of the Green's function or a sum over states [1]. More recently, real-space approaches have also been considered [2–4]. Alternative methods based on many-body perturbation theory have also been suggested [5,6]. However, all of these approaches scale with system size N roughly as N^4 [7]. Imaginary-time techniques have been shown to reduce the scaling to N^3 [8–10]. Here we present improvements to the localized, real-space approach originally introduced in Ref. [11]. We revisit the local, real-space approach for two broad reasons: improvements of the accuracy of the calculations and improvements to the execution of the code, yielding a scaling of $N^2 \log N$. Highly accurate calcu-

lations are necessary to understand subtle changes in spectral features due to changes in crystal or electronic structure. Improved algorithmic scaling and an efficient implementation are necessary to make screening calculations practical for large systems.

In the first part of this paper, we extend the calculation of the polarizability to include all-electron projectors and examine the effect of the pseudopotential approximation on the screened Coulomb potential. We find that pseudopotential-based wave functions yield an incorrect polarizability for small distances, e.g., inside the pseudopotential radius. The removal of the core-level electrons removes nodes from the valence atomic orbitals, changing the density distribution of the valence orbitals. The discrepancy is small but has a noticeable effect on calculated near-edge x-ray absorption spectra, where the strength of excitonic binding is highly dependent on the screening provided by the valence electrons. We have extended this screening approach for use in valence excitations. In the case of valence excitations (UV/optical spectra), we find that augmentation is not necessary.

Second, we present a substantial improvement to the system-size scaling of RPA calculations. Our method scales as $N^2 \log N$, while still performing well for small systems sizes. It also does not require the dense k-point grids of traditional reciprocal-space methods. Even for small unit cells, a 2^3 k-point grid can be sufficient. This not only reduces the computational cost of the screening calculation itself, but also the cost of generating electron wave functions on a dense k-point grid. The calculation of the polarizability has been implemented as a hybrid OPENMP+MPI code, allowing screening calculations of large systems to be carried out quickly. The local, real-space approach also provides an ideal testbed for investigating higher-order screening methods.

In Sec. II, we review polarizability and screening and the main approximations of the real-space approach. In Sec. III,

*john.vinson@nist.gov

we detail how the use of pseudopotentials affects the orbitals near an atom which in turn modifies the electronic screening. We introduce the optimal projector functions which are used to augment the pseudopotential-based orbitals, restoring their all-electron character, and we examine the effect this has on calculations of the screened Coulomb potential. In Sec. IV, we illustrate the effect of augmentation on Bethe-Salpeter equation calculations of x-ray absorption through modifications of the core-hole screening. In Sec. V, we generalize the local, real-space approach for use in calculating valence (optical/UV) excitation spectra, and we show that augmentation is unnecessary for valence calculations. In Sec. VI, we detail the performance of our implementation, demonstrating the superior system-size scaling of our method. Last, in Sec. VII we summarize future enhancements and applications for our method, including an extension to beyond-RPA screening.

II. REVIEW

A. Polarizability and screening

We begin by reviewing the definitions of the polarizability and the dielectric response that screens an applied, external potential. We use Hartree atomic units throughout, such that the electron charge, electron mass, Planck's constant and Coulomb's constant are given by $e = m_e = \hbar = 4\pi\epsilon_0 = 1$. We make use of the one-electron Green's function

$$g^{-1}(1, 2) = g_0^{-1}(1, 2) - V(1, 2) - \Sigma(1, 2), \qquad (1)$$

where each numerical index denotes position, time, and spin; g_0 is the noninteracting Green's function; V is the total potential, which is local and so contains a factor $\delta(1 - 2)$; and Σ is the self-energy encompassing many-body interactions.

The irreducible polarizability χ_0 of the electron system is the change in electron density n in response to a change in the total potential V,

$$\chi_0(1, 2) = \delta n(1)/\delta V(2). \qquad (2)$$

The density can be written in terms of the one-electron Green's function,

$$n(1) = -i g(1, 1^+), \qquad (3)$$

where 1^+ refers to a time infinitesimally later than 1. Taking functional derivatives with respect to the potential in Eq. (1), the polarizability can be written

$$\chi_0(1, 2) = -i \delta g(1, 1^+)/\delta V(2)$$
$$= \int d3 d4 \, i g(1, 3) \frac{\delta g^{-1}(3, 4)}{\delta V(2)} g(4, 1^+). \qquad (4)$$

By approximating the Green's function using $g^{-1} = g_0^{-1} - V$, we arrive at the random-phase approximation for the polarizability,

$$\chi_0^{RPA}(1, 2) = -i g_0(1, 2) g_0(2, 1^+). \qquad (5)$$

This can be transformed from the two-time representation to the response as a function of a single external energy ω and written in real space: $\chi_0(\mathbf{r}, \mathbf{r}', \omega)$.

Above, in Eq. (1), the potential term includes both the external and Hartree terms, and the Hartree term itself will change with changes in the electron density. One should therefore use the reducible polarizability χ which is the response to only changes in the external potential,

$$\chi(\mathbf{r}, \mathbf{r}', \omega) = \chi_0(\mathbf{r}, \mathbf{r}', \omega)$$
$$+ \chi_0(\mathbf{r}, \mathbf{x}, \omega) v(\mathbf{x}, \mathbf{x}') \chi(\mathbf{x}', \mathbf{r}', \omega), \qquad (6)$$

where repeated spatial indices $\mathbf{x}$ are integrated over. Here v is the Coulomb operator. Most importantly, from the reducible polarizability one obtains the screened Coulomb operator $W(\mathbf{r}, \mathbf{r}', \omega)$

$$W(\mathbf{r}, \mathbf{r}', \omega) = \epsilon^{-1}(\mathbf{r}, \mathbf{x}, \omega) v_{\text{ext}}(\mathbf{x}, \mathbf{r}')$$
$$= v_{\text{ext}}(\mathbf{r}, \mathbf{r}') + v(\mathbf{r}, \mathbf{x}) \chi(\mathbf{x}, \mathbf{x}', \omega) v_{\text{ext}}(\mathbf{x}', \mathbf{r}'), \qquad (7)$$

where ϵ is the dielectric tensor. This is central to many-body perturbation techniques, for instance, for treating single-particle self-energies via the GW method or electron-hole excitation calculations via the Bethe-Salpeter equation (BSE) [12].

B. Real-space decomposition of the screening

One can divide a potential into pieces and separately consider the screening of each piece. The two-coordinate external Coulomb operator in Eq. (7) can, without loss of generality, be written with a strength q and parametric dependance on the second spatial coordinate

$$v_{\text{ext}}^{[\mathbf{r}']}(\mathbf{x}) = q|\mathbf{x}|^{-1} = q|\mathbf{r} - \mathbf{r}'|^{-1}, \qquad (8)$$

where $\mathbf{x} = \mathbf{r} - \mathbf{r}'$. Following Ref. [11], this potential is partitioned by adding and subtracting a shell of charge:

$$v_{\text{ext}}(\mathbf{x}) = v_1(\mathbf{x}) + v_2(\mathbf{x}),$$
$$v_1(\mathbf{x}) = v_{\text{ext}}(\mathbf{x}) - v_2(\mathbf{x}),$$
$$v_2(\mathbf{x}) = q R_S^{-1} \Theta[R_S - x] + q x^{-1} \Theta[x - R_S]. \qquad (9)$$

Here Θ is the Heaviside theta function and R_S is the shell radius. The screened potential is therefore

$$W^{[\mathbf{r}']}(\mathbf{r}, \omega) = \epsilon^{-1}(\mathbf{r}, \mathbf{x} = \mathbf{r} - \mathbf{r}', \omega) v_{\text{ext}}^{[\mathbf{r}']}(\mathbf{x})$$
$$= \epsilon^{-1} v_1 + \epsilon^{-1} v_2$$
$$= W^{(1)} + W^{(2)}. \qquad (10)$$

To this point no approximation has been made in the screening of v_{ext}. The full, two-coordinate $W(\mathbf{r}, \mathbf{r}')$ is recovered by evaluating the decomposed, real-space screening at each $\mathbf{r}'$, as is shown for valence BSE calculations in Sec. V.

C. Local, real-space approach

The utility of the local, real-space approach rests on treating the screening of v_1 and v_2 differently. Because v_1 is nonzero only within R_S, the dielectric response that screens v_1 is also localized. By only calculating the polarizability within a finite volume, we are able to reduce the computational cost of the calculation. In contrast, the screening of v_2 must still cover all space. We therefore treat the screening of v_2 only approximately, using a model for the dielectric response.

Our method is therefore an approximation to the RPA screening, controlled by the shell radius R_S and the model screening used to screen v_2. As R_S increases, v_2 goes to zero, and treating v_2 approximately is a *controlled* approximation. The effect of dividing the calculation at finite R_S is addressed further in Sec. IV C and Appendix D 3. A modified Levine-Louie dielectric model was adopted in Ref. [11] and is used here [13,14]. This model screening is parametrized by the static, long-range dielectric constant ϵ_∞ and the average and local valence electron densities.

This choice makes ϵ_∞ an input parameter, but small errors in $1/\epsilon_\infty$ will only lead to small errors in the resulting x-ray absorption spectra, and we investigate this further in Sec. IV C and Appendix D 4. As suggested in Ref. [11], the preferred source is an experimental measurement of the optical constants or index of refraction below the band gap. If such data are not available, several computational approaches can be used. The optical spectra can be calculated with the OCEAN code [15,16], either within the RPA or within the BSE approximation, which formally requires iteration to equate the input and output ϵ_∞.

The short-ranged potential v_1 is screened by calculating the RPA polarizability. Typically we use a shell radius R_S between 3 and 5 a.u. The polarizability is calculated within a spherical region of space given by a radius of 8 to 10 a.u. Per Eq. (5), this necessitates calculating the one-electron Green's functions for this region of space. Improvements in the fidelity, efficiency, and parallelization of calculating g, χ_0, and χ in order to screen v_1 are the focus of this work.

III. AUGMENTATION OF ELECTRON ORBITALS

A. Pseudopotentials

The screening method of Ref. [11], adopted in the OCEAN spectroscopy code, utilizes electron wave functions generated from a pseudopotential-based density-functional theory code. The pseudopotential approximation allows for a dramatic reduction in the computational cost of plane-wave codes. The core-level electrons are removed, and the $-Z/r$ potential of each ion is replaced by a pseudopotential, such that the valence electrons are the most-bound states. This reduces the computational cost by treating fewer electrons and smoothing the remaining electrons wave functions, reducing the required number of plane waves. More information on pseudopotential theory and methods can be found in Ref. [17].

Consider a pseudopotential $V_{\text{ion}}^{\text{ps}}$ for an element which treat angular-momentum-dependent effects separably following the Kleinman-Bylander form [18]. For each value of the principal angular momentum l up to l_{max} it consists of some local potential and optionally some nonlocal projectors. In contrast, the all-electron ionic potential is simply an attractive Coulomb potential set by the ion's atomic number Z. Both the all-electron (ae) and pseudized (ps) radial Schrödinger equations can be solved numerically,

$$H^{\text{ae/ps}}\phi_j^{\text{ae/ps}} = \varepsilon_j \phi_j^{\text{ae/ps}}$$

$$H^{\text{ae/ps}} = -\frac{\nabla^2}{2} + \frac{l(l+1)}{2r^2} + V_H[n] + V_{\text{xc}}[n] + V_{\text{ion}}^{\text{ae/ps}}.$$

(11)

The Hartree V_H and density-functional theory (DFT) exchange-correlation V_{xc} potentials both depend on the density n, and so the problem must be solved self-consistently. Outside of some cutoff radius r_c the pseudopotential matches the all-electron ionic potential, and the pseudo and all-electron electron orbitals are equal for an isolated atom. Additional requirements are enforced at r_c. The all-electron and pseudo orbitals should have the same radial derivative and scattering length, although this is exact only for specific energies. Because the scattering properties are the same for the all-electron and pseudo wave functions, the behavior of the pseudo electrons in the interstitial regions is identical to that of the electrons in the all-electron system.

While pseudopotentials are capable of reproducing all-electron results for many properties, such as band structures (Bloch-state energies) or structural properties, they fall short in others, such as core-level excitations. This is unsurprising because there are no core-level electrons in the pseudopotential system. We point to localization and location of a perturbation or transition operator in determining if the pseudopotential wave functions are sufficient. In the case of core-level transitions, the operator is both highly localized and within the pseudized region. Conversely, structural properties are not localized. They rely on force constants, which are a measure of the change in the distribution electron density as felt by one atom in response to the motion of another. In the case of a real-space screening approach, we are interested in the density response to a localized perturbation, and for screening a core hole this localized perturbation is within the pseudized region.

To demonstrate the differences between an all-electron and pseudopotential screened core-hole, in Fig. 1(a) we show the difference in the valence ($2s$ and $2p$) electron densities between the all-electron and the pseudopotential in a fluorine atom. While the all-electron $2s$ orbital has a node (and is orthogonal to the $1s$ orbital), the pseudopotential $2s$ does not. The first antinode of the all-electron $2s$ is responsible for the difference in densities below around 0.2 a.u. Around 0.7 a.u. the pseudopotential density exceeds the all-electron density, a consequence of norm conservation and the deficiency at small radii. In Fig. 1(b) we plot the change in valence electron density on introduction of a $1s$ core hole in the system. For both systems the valence electron density moves to smaller radii (towards the positively charged core-hole), but for the all-electron system there is a larger change at short distances.

The small difference in density response is magnified when converted into an induced potential as shown in Fig. 1(c). Clearly, at short distances the all-electron valence orbitals are more efficient at screening core-level excitations than the pseudopotential orbitals. In Sec. IV we will explore how this difference in induced potential affects calculated x-ray absorption spectra that can be compared to measured ones.

In the context of screening a core hole, the potential v_{ext} is already weakened by relaxation of the core orbitals that are not included explicitly when valence screening effects are computed. This largely prevents problems that could arise from a frozen-core approximation regarding changes in the core level occupancies. Smaller, additional core-relaxation effects arise when there are changes in the chemical environment. These can include core polarization (mostly dipolar)
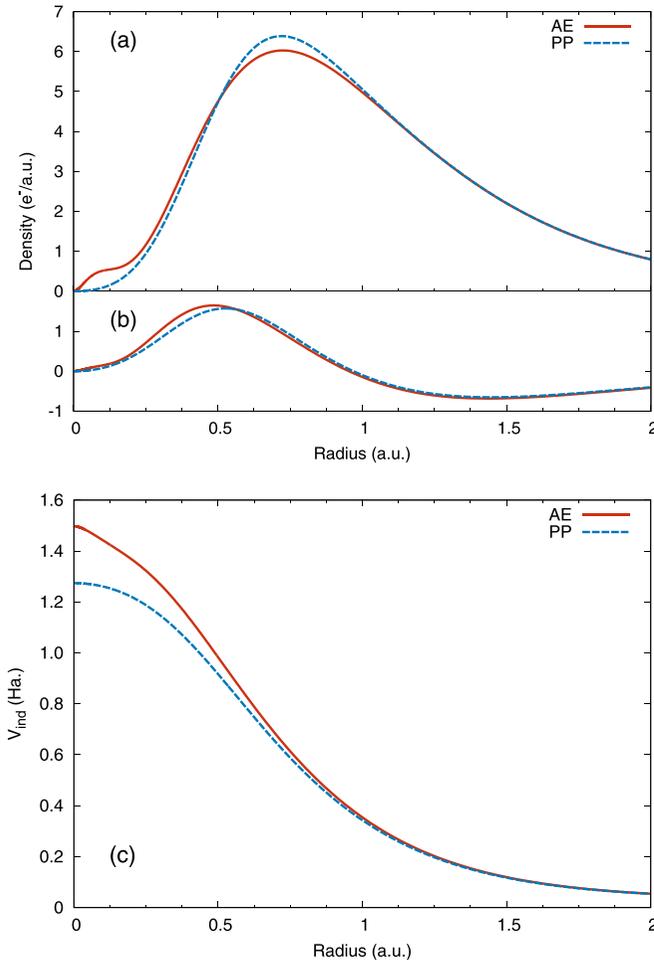
FIG. 1. (a) Valence electron density for an isolated fluorine atom for both the all-electron (AE, solid red) and pseudopotential (PP, dashed blue) systems. In the pseudopotential system the 2*s* orbitals have no nodes. This is why the all-electron systems has more density at very small radii, inside the 2*s* node at 0.2 a.u. (b) The change in valence density of the fluorine atom in response to a 1*s* core hole. (c) The induced potential that arises from the change in the valence density. Even though the difference in densities is quite small, the subsequent difference in the induced potential is substantial at short ranges.

and relaxational self-energies of valence electrons or holes added onto an atomic site. One method to treat many of these effects is presented by Shirley and Martin [19], who also refer to previous work.

B. Optimal projector functions

To include the correct all-electron behavior, we augment wave functions of the pseudopotential system. This augmentation style was introduced for core-level transitions [15,20] and borrows heavily from the projector-augmented wave (PAW) method [21]. In the PAW method, Bloch states are augmented by projecting smoothened wave functions onto a basis set of projector functions. For DFT ground-state calculations, it may be advantageous to optimize the projection scheme for describing the highest occupied and lowest unoccupied states, whereas our emphasis is obtaining realistic electron wave functions over a wide range of occupied and unoccupied bands. The coefficient of each projector function weighs the correction of the wave function in the form of replacing the projector function with an all-electron counterpart.

In OCEAN, all-electron and pseudo versions of partial waves are evaluated and condensed into a set of optical projector functions (OPFs). Typically, partial waves are sampled at several dozen regularly spaced energies over a multi-Hartree energy range. The partial waves are constructed within an augmentation radius the encompasses the the pseudotpotential cut-off radius $r_a \geqslant r_c$. The OPFs are the eigenvectors that have the largest eigenvalues of the overlap matrix between all of the partial waves, considering each angular momentum separately. As a result, the OPFs are not single-energy partial waves, in contrast to typical PAW construction. The eigenvalues fall off rapidly enough that over 99.9% of the wave function's degrees of freedom are sampled with only a few OPFs. This is described in further detail in Appendix C.

Augmentation exploits the same properties that we earlier asserted our pseudopotential would have. Namely, outside of the pseudopotential cut-off radius r_c the pseudo wave function is identical to the all-electron wave function. Over a reasonable energy range, a bound or scattering state can be transformed from the pseudopotential system to the all-electron system by replacing the pseudo wave function inside of the cut-off radius with the all-electron solution. The OPF method is only used for augmentation to reconstruct wave functions as a means of postprocessing DFT results, and the implementation is discussed in Sec. VI C.

C. Approximate augmentation

Previously, the OCEAN code relied on an approximation instead of carrying out augmentation of the wave functions. We document the old approach here, and in the next section we will compare it to the current method. As was shown in Fig. 1(c), even the isolated atom demonstrates the importance of all-electron wave functions when calculating the screening near the nucleus. The difference between the two induced potentials in Fig. 1(c) can be calculated purely within the isolated atomic case

$$\Delta v_{\mathrm{ind}}(r) = v_{\mathrm{ind}}^{\mathrm{ae}}(r) - v_{\mathrm{ind}}^{\mathrm{ps}}(r). \tag{12}$$

Previously this correction was at times applied to the induced potential as calculated within the RPA using the *unaugmented* wave functions of the system. In this way the approximate effect of augmentation was included in the screening. However, because this method relies on calculations of an isolated atom, the valence orbitals and their occupations are only a rough approximation to the system of interest.

D. Screening with augmented orbitals

Before showing the effects of augmentation on calculated spectra, we examine the effects on the screening, using LiF as a model system. In Fig. 2 we show the spherically averaged induced potential in response to a F 1*s* core hole. As in the case of the isolated atom (Fig. 1), the induced potential is stronger near the origin when the orbitals have all-electron character, and the effect of augmentation is only significant near the
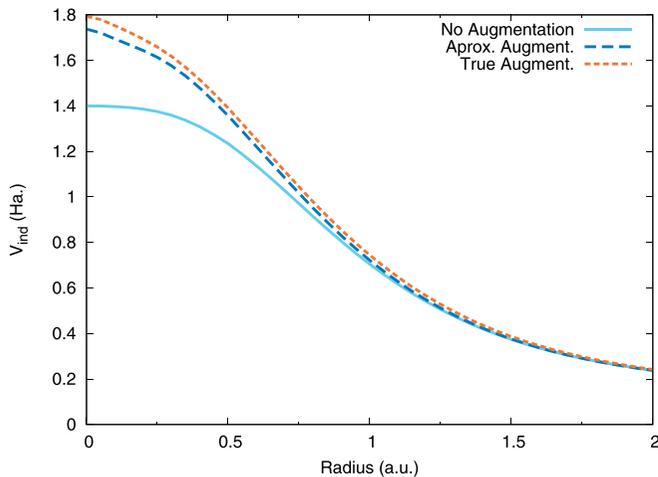
FIG. 2. The spherically averaged induced potential in response to a F $1s$ hole in LiF. The same calculation is carried out three times using different treatments for correcting the electron orbitals near the F atom (see text). The difference between the no augmentation and true augmentation potentials has the same origin as in Fig. 1(c) for the case of an isolated F atom.

atom where the all-electron and pseudopotential orbitals are different. The approximate augmentation, while exact for an isolated atom, does not fully reproduce the screening using augmented orbitals within a solid. It is also only valid for atom-centered response such as for core-level spectroscopy.

For valence-level spectroscopy or other calculations, such as self-energy calculations using the GW method, the screening must be calculated throughout the unit cell. In Fig. 3 we show the induced potential in response to test charges centered along the $(a/2, a/2, a/2)$ (from F to Li) and $(0, a/4, a/4)$ (from F halfway to the nearest F) directions within the unit cell, where $a = 4.0$ Å. Unsurprisingly, the effect of augmentation is only noticeable near the F atom. Note that while the point $(a/2, a/2, a/2)$ is centered on a Li atom, there is no effect from augmentation because the Li $1s$ orbitals are included as valence.

IV. CORE-LEVEL SPECTROSCOPY

We now will use calculations of x-ray absorption spectra (XAS) to investigate the effect of augmenting the orbitals and the robustness of the approximations in real-space method. Starting with LiF, we show the importance of the screened electron–core-hole interaction and demonstrate that the effect of augmentation is comparable to the changes observed due to thermal disorder. Using the series of lithium halides, we show that the approximate augmentation method is sufficient for heavier ions and that the importance of augmentation is reduced by including semicore orbitals in the pseudopotential. Finally, using hexagonal boron nitride, we show that the real-space method is broadly applicable, including for layered materials with high levels of site anisotropy.

Within the BSE approach, absorption spectra are modeled by considering an interacting electron-hole pair [12,15,16]. The strong Coulomb attraction between the electron and the core hole is screened by the dielectric response of the material.
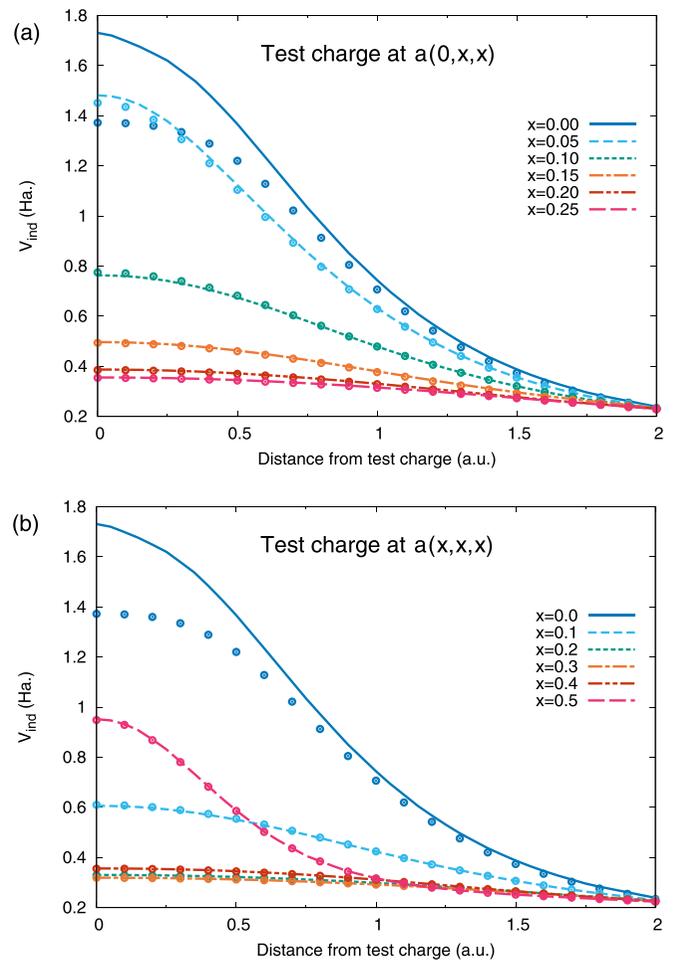

FIG. 3. (a) The spherically averaged induced potentials in response to test charges in LiF as a function of the distance from the test charge. The potential is shown calculated both with augmentation (lines) and without (circles). The test charges are centered at points along the line segment from the F atom at the origin to halfway to the nearest-neighbor F atom $(0, a/2, a/2)$, where $a = 4.0$ Å is the lattice constant. Only near the F atom are the effects of augmentation observable. (b) The same, but along (a, a, a). The Li atom is located at $x = 0.5$ [i.e., $(a/2, a/2, a/2)$]. Because the $1s$ orbitals are included as valence there is no effect of augmentation around it.

The calculated XAS depends strongly on the strength of this attraction, and, therefore, the details of the screening. Within OCEAN, the BSE calculations are carried out using a basis of electron orbitals calculated within DFT. Here we use the QUANTUM ESPRESSO code [22,23] and the local-density approximation for the density functional [24]. Pseudopotentials are taken from PseudoDojo [25] and generated with ONCVPSP [26]. Various convergence parameters are summarized in Appendix A.

A. Effect of augmentation on XAS calculations

We first consider the fluorine K edge in LiF, which has been studied with OCEAN previously [15,27]. Figure 4 shows the effect of changes to the screened Coulomb attraction between the electron and core hole. In light gray, the non-interacting spectrum (neglecting electron-hole interactions) is
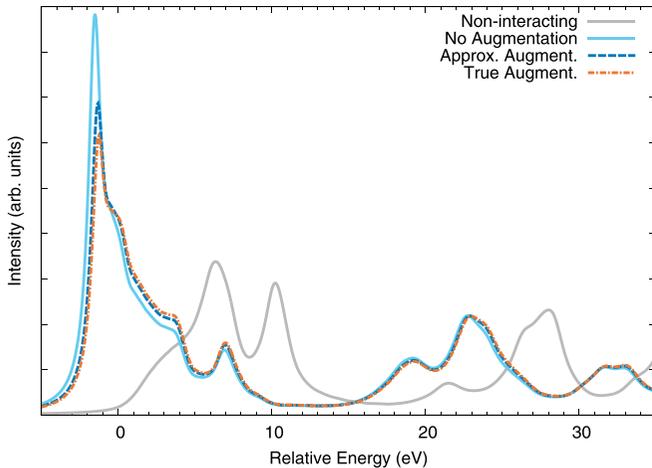
FIG. 4. The fluorine K-edge absorption in LiF using different approximations for the screened Coulomb interaction. Augmentation of the electron wave functions leads to a stronger screening of the core-hole potential. This stronger screening reduces the strength of the exciton and blue-shifts the spectra.
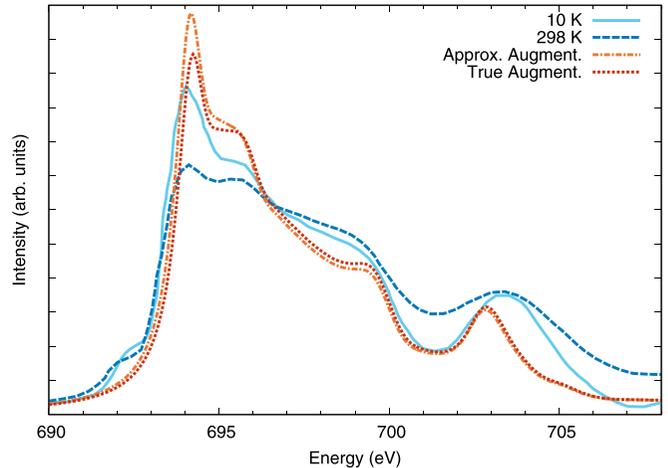

FIG. 5. The fluorine K-edge near-edge absorption in LiF comparing calculations using approximate (orange, dot-dash) and true (red, dot) augmentation for the screening for measurements at 10 K (light blue, solid) and 298 K (dark blue, dashed). The difference in strength of the main exciton at 694 eV between the two calculations is comparable to the difference seen in the two measurements due to increased vibrational disorder at room temperature. Measured data were taken from reference [27].

dramatically different from any other approximation, and it shows the importance of the excitonic binding on the absorption spectrum. The three different approximations to the screening all give qualitatively the same results, capturing a strong exciton around 1.5 eV below the onset of the non-interacting spectrum (0 eV in the plot). The differences are mostly confined to the near-edge region, within 10 eV of the onset. Without any augmentation the exciton is 1.52 eV below the conduction band minimum. This reduces to 1.34 eV with approximate augmentation and to 1.29 eV with true augmentation. Correspondingly, the exciton strength is reduced with augmentation, improving agreement with experiment (Fig. 5).

Small changes in intensity and spacing of peaks in the x-ray absorption spectra can be signs of changes in the local structure around the absorbing atom [28]. To illustrate the importance of accurately determining the relative intensities of peaks, we reproduce measured XAS from reference [27] which shows the change in the F K-edge absorption with temperature. LiF has a dipole-forbidden pre-edge that is observable in dipole-limited absorption spectra due to the vibration of the ions which breaks the symmetry around the F atoms [27,28]. When simulating spectroscopy of condensed systems, disorder is typically treated within the Born-Oppenheimer approximation where the ions are stationary during the excitation process. The final calculated spectrum is an average over spectra from different atomic positions, e.g., generated using molecular dynamics simulations [28] or statistical sampling of the vibrational modes of the system [29], or a statistical average of core-hole displacements [30]. Typically in calculations of extended system, the vibrational energy levels of the ionic system are treated as negligible and Frank-Condon effects are neglected [30].

In Fig. 5 we compare the calculated changes in the F K near-edge spectra of LiF due to changes in the screening calculation to measured changes due to temperature. At increased temperature, the pre-edge feature near 692 eV moves

to lower energy, all of the features are broadened, and there is a noticeable weakening of the main exciton between 694 and 696 eV. The main exciton also differs in strength between the calculations using different augmentation methods, which is even more exaggerated if augmentation is neglected (Fig 4). While the error in the calculation from neglecting or approximating the augmentation has only a small qualitative effect on the spectrum, the differences are substantial when compared to small structural changes. High-fidelity screening calculations are necessary for correctly identifying local structure or assessing the accuracy of approximations used for incorporating disorder or vibrations.

B. Influence of valence principal quantum number and semicore orbitals

Having shown the large effect of changes to the screening on the XAS of LiF, we next examine the effect of augmentation on heavier ions by exploring the range of lithium halides. All crystallize in the rocksalt $Fm\bar{3}m$ structure, and for all four materials a uniform core-hole lifetime broadening of 0.5 eV was used. In Fig. 6 we show the effects of different augmentation approaches on the halide K edges of LiCl, LiBr, and LiI. The trend between approximations for the same compound shown for LiF in Fig. 4 holds for the heavier halides as well. Calculating the screening of the core-hole potential with wave functions from a pseudopotential calculation dramatically under-screens the core hole, resulting in an exaggerated excitonic peak. There is a trend towards smaller discrepancy with increasing atomic number in the halide series. For the LiCl the exciton is approximately 0.56 eV (0.26 eV) more bound without any augmentation (with approximate augmentation), while for LiBr the overbinding is 0.65 eV (0.11 eV) and 0.05 eV (0.01 eV) for LiI. We conclude that some all-electron augmentation is necessary for the proper calculation
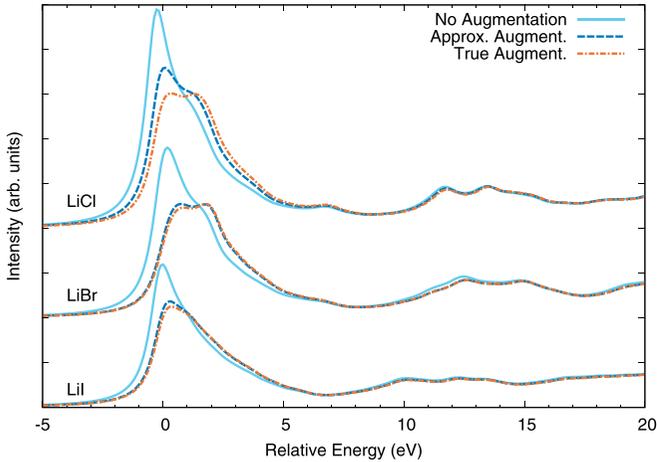
FIG. 6. The halide K-edge absorption for LiCl, LiBr, and LiI using different approximations for the screened Coulomb interaction. Across the three materials proper augmentation leads to increased short-range screening and weaker excitonic peaks.

FIG. 7. The iodine K-edge absorption for LiI using either the standard (Std.) iodine pseudopential ($4d5s5p$) or a semicore (S-C) iodine pseudopotential ($4s4p4d5s5p$) and either no augmentation in the screening or true augmentation. The presence of semicore orbitals substantially reduces the error if no augmentation is used. With augmentation the two pseudopotentials give equivalent spectra.

of the screening even for heavier atoms, but for Br and I, it appears that the approximate augmentation method may be sufficient.

The differences are primarily confined to the near-edge region, which could be expected from Fig. 1(c). As for LiF, the differences between the augmented and unaugmented induced potential are confined to a small region around the core hole. Only near the edge onset is the excited electron localized enough to be strongly affected by this very localized difference in core-hole potential. However, an increase in spectral weight near threshold necessitates a reduction in spectral weight higher in energy—high-energy states in the noninteracting system are pulled down by the core hole. The differences shown here would be mitigated somewhat by more realistic core-hole lifetime broadening for the Br and I K edges, though by observing partial fluorescence emission from the $2p_{3/2}$, the 0.5 eV broadening used here remains realistic for the Br.

Next, we examine the effects of augmentation by comparing calculations of LiI using two different iodine pseudopotentials. The standard iodine pseudopotentials uses a Kr core with the $4d$, $5s$, and $5p$ electrons in the valence bands while the semicore pseudopotential also includes the $4s$ and $4p$ orbitals as valence. In Fig. 7 we show that the calculated I K edge of LiI does not depend on the pseudopotential when the orbitals are properly augmented for the screening calculation. However, without augmentation the screening calculated using the standard pseudopotential is notably weaker than that of the semicore pseudopotential, leading to a stronger exciton. The weaker effect of augmentation on the calculation using the semicore pseudopotential is due to the node in the $5s$ and $5p$ orbitals (absent for the standard pseudopotential). As discussed earlier (see Fig. 1), the absence of nodes in the pseudopotential orbitals shifts the electron density away from the atom. Even though without augmentation the $n = 5$ orbitals only get a single node in the semicore system the effect is already dramatic and the discrepancy due to neglecting augmentation is strongly reduced.
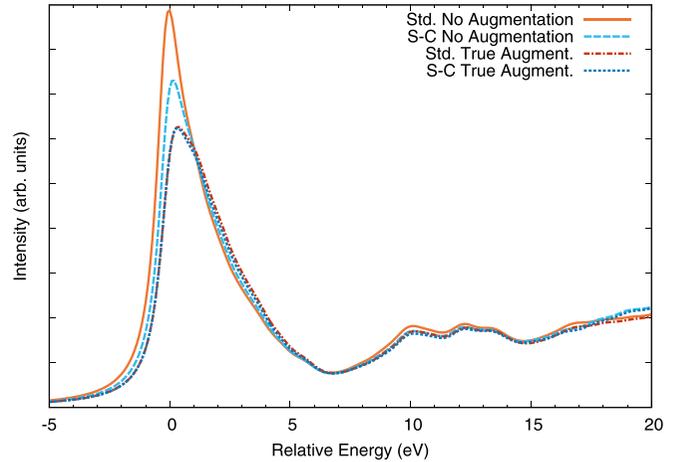
We have demonstrated that the pseudopotential-based orbitals are insufficiently accurate for calculating the screened electron-hole interaction for core-level spectra. This failure is independent of the quality of the pseudopotential used. Instead it is a straightforward consequence of the difference in nodal structure between valence orbitals in the all-electron and pseudopotential systems. As shown in the atomic case (Fig. 1), the absence of nodes in valence orbitals of the pseudopotential system affects the density response and hence the screening. This is mitigated somewhat by the use of semicore pseudopotentials that include the next highest principal quantum number in the valence, e.g., Fig. 7. The moderate success of including semicore orbitals without augmentation may indicate that a kind of false augmentation could be used, where nodes are added without requiring that they mimic the true all-electron orbitals. These false projectors could be constructed such that they add nodes without increasing the required plane-wave energy cutoff for use in reciprocal-space methods.

C. Layered materials, convergence, and errors

We next investigate the K edges of hexagonal boron nitride (*h*-BN). The XAS of *h*-BN has been the subject of recent investigations into the role that vibrational disorder and defects play in the spectra [31–33]. Because *h*-BN is layered, it has an anisotropic dielectric response, unlike the lithium halides. We will use it to showcase the real-space screening method on systems with reduced symmetry. The dielectric screening within the BN planes (perpendicular to the *c* axis) is stronger than that parallel to the *c* axis: $\epsilon_\infty^\perp = 4.95$ versus $\epsilon_\infty^\parallel = 4.10$ [34]. The anisotropy is also visible in the XAS. Both the B and N edges show strong excitonic features when the x-ray polarization vector is aligned with the *c* axis and a delayed onset when the x-ray polarization is in-plane.

We compare our calculations of *h*-BN to two different computational approaches. First, the EXCITING code (like OCEAN)
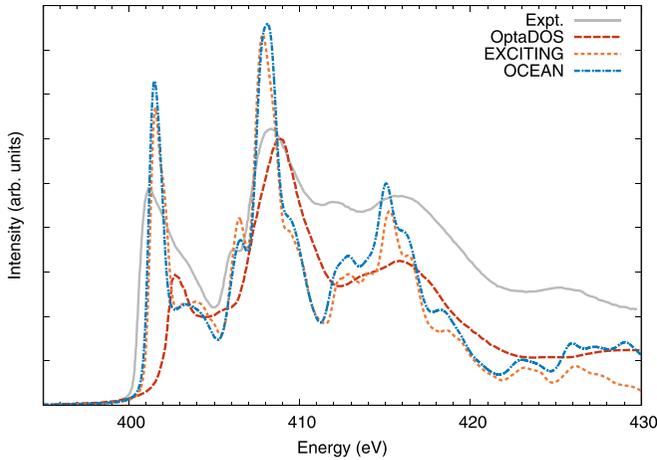
FIG. 8. The measured nitrogen K-edge x-ray absorption of h-BN compared to three different computational methods. The two BSE-based codes, EXCITING and OCEAN, show good agreement with each other and experiment. The experiment is reproduced from Ref. [39] and the OptaDOS from Ref. [40].

uses the BSE [35,36]. The screening in EXCITING is also carried out in the RPA, but calculated in reciprocal space with no modeled response. Additionally, it is an all-electron code making augmentation unnecessary. Second, within OptaDOS x-ray spectra are calculated using the ΔSCF method [37]. In this approach the final states are calculated directly as the unoccupied states of a DFT calculation with a core hole. The initial state is a the core-level orbital, and the transition matrix elements can be calculated directly. The density response to the core hole is calculated self-consistently within DFT. The OptaDOS calculations reproduced here used the CASTEP DFT code [38].

In Fig. 8 we compare the N K-edge of h-BN calculated with OCEAN to experiment [39] and calculations using OptaDOS and EXCITING. The experiment was taken with polarization at an angle of $50°$ to the c axis, giving nearly the same 2:1 in-plane to out-of-plane ratio of a disordered sample. All three calculations are averaged over polarization directions. The ΔSCF OptaDOS spectrum is a clear outlier with a substantially reduced exciton strength and binding at the onset, a known characteristic of ΔSCF methods [41]. The agreement between OCEAN and EXCITING is very good. The EXCITING calculation includes fewer conduction bands, leading to an artificial die-off in spectral intensity at higher energies.

Per Sec. II C, the real-space screening method is only an approximation to the RPA response. The approximation is controlled by R_S, the radius of the shell of charge that is screened with a model dielectric response, and ϵ_∞, the value of the static dielectric constant which is input into the Levine-Louie dielectric model (see Appendix B). In Fig. 9(a) we show the total induced potential in h-BN due to a core hole on the nitrogen atom. We also show the difference in the potential calculated using a shell radius of 7 a.u. versus smaller radii, and we find that the errors are less than 3 mHa. Next we vary the value of the input dielectric constant by approximately $\pm 15\%$, plotting the difference in the potentials in Fig. 9(b).
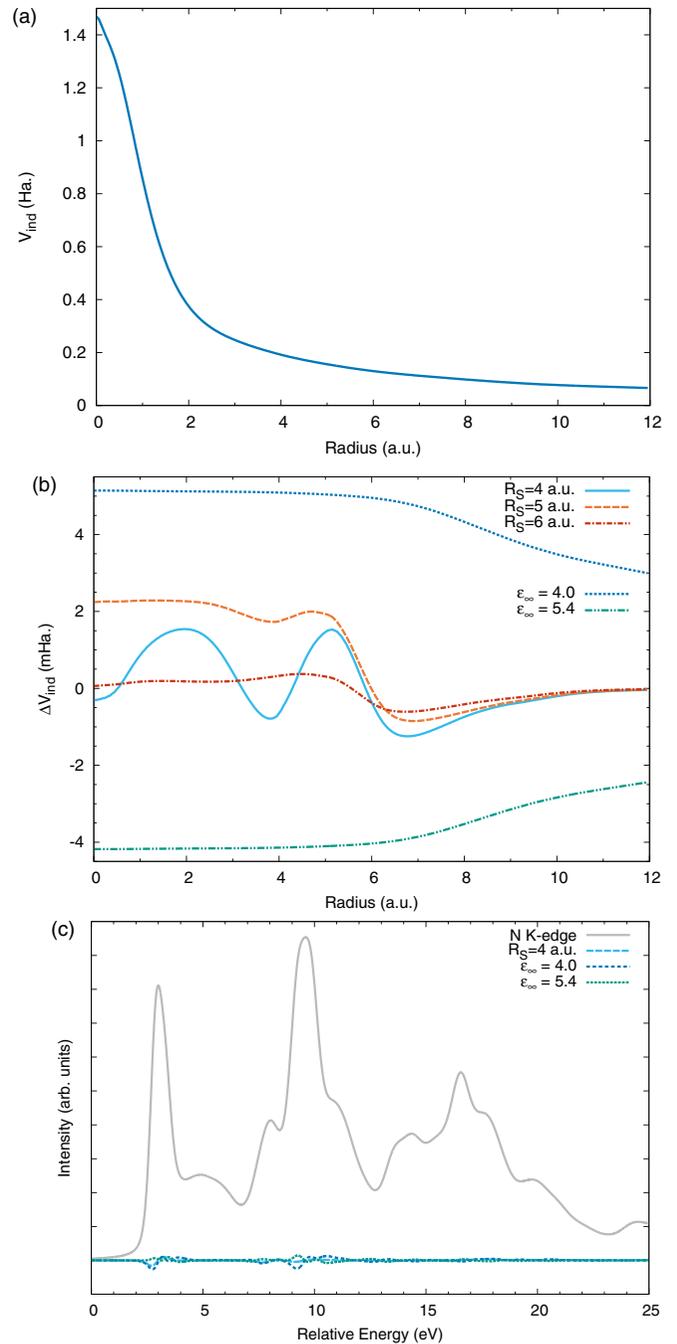

FIG. 9. (a) The induced potential due to a $1s$ core hole on a nitrogen atom in h-BN as a function of radius, using a spherical charge radius of $R_S = 7$ a.u. and a dielectric constant of $\epsilon_\infty = 4.667$. (b) The differences in the calculated induced potential due to changes in either the shell radius or input dielectric constant. (c) The calculated N K-edge XAS of h-BN (gray, solid). The differences between the spectrum calculated from the converged screening and the spectra using different changes to the screening calculation. Before taking the differences, the XAS calculations are aligned to correct for red or blue shifts (see text).

As expected, increasing the dielectric constant input into the model decreases the induced potential. However, we find the differences are $\lesssim 5$ mHa. The screening calculation is relatively insensitive to changes in the sphere radius or errors in

the input dielectric constant. Further examination is presented in Appendix D.

The primary effect of small errors in the screening is a shift of the spectra, reflecting a change in the excitonic binding energy. For the nitrogen K edge, using $\epsilon_\infty = 5.4$ blue-shifts the x-ray absorption by 0.10 eV, while using $\epsilon_\infty = 4.0$ red-shifts the x-ray absorption by 0.12 eV. In Fig. 9(c) we show how the nitrogen K-edge spectra change with these changes in the calculation of the screening. The changes are minor, and so only the differences between the spectra are plotted, after taking into account the aforementioned red or blue shifts.

V. VALENCE EXCITATION SPECTROSCOPY

The OCEAN code is also capable of calculating valence optical/UV spectroscopy within the BSE, following earlier work [42,43]. In moving from core to valence excitations, the hole is no longer confined to a local basis around the atom (the core-level orbital), but instead spans the unit cell. This introduces a trade-off. For a valence calculation, the screening must be calculated throughout the unit cell, but the more delocalized nature of a valence exciton means that the importance of the screening calculation at any given point in space is diminished. Previous valence calculations have forgone RPA calculations and used a model dielectric function while still achieving good agreement with experiment [42,43]. Conversely, using only a model dielectric to screen core-level excitations had been found to be inaccurate [14]. The strong colocality of the core hole and photoelectron in near-edge x-ray absorption makes the spectra very sensitive to the accuracy of the screened Coulomb potential at small distances where the accuracy of an electron-gas-based model model breaks down.

A. Screening for valence-level BSE

Previously in OCEAN, the screened potential for valence calculations was approximated using the Hybertsen-Levine-Louie dielectric model [13,44], which depends parametrically on the local density $\rho(\mathbf{x})$ and static dielectric constant ϵ_∞. The wave functions for the electrons and holes are sampled on regular real-space grids $\mathbf{x}$, and therefore we need a description of the screened Coulomb potential for each set of grid points $W(\mathbf{x}, \mathbf{x}')$. Following Ref. [42], the screened Coulomb potential is given by

$$W_{\mathrm{HLL}}(\mathbf{x}, \mathbf{x}') = {}^{1}/_{2}[W_{\mathrm{hom}}(|\mathbf{x} - \mathbf{x}'|; \rho(\mathbf{x}), \epsilon_\infty)$$
$$W_{\mathrm{hom}}(|\mathbf{x} - \mathbf{x}'|; \rho(\mathbf{x}'), \epsilon_\infty)], \qquad (13)$$

which simply averages the results using the density at points $\mathbf{x}$ and $\mathbf{x}'$. To avoid the divergence at $\mathbf{x} \to \mathbf{x}'$, a spherical average over the discretization volume is used when $\mathbf{x} = \mathbf{x}'$.

To improve this, we substitute the more accurate local-RPA result for the short-range part of W. Using the previously introduced method, we can calculate the screened Coulomb from Eq. (10) for each grid point $\mathbf{x}$, ie, $W^{[\mathbf{x}]}$. In the case of core-level excitations the external potential in Eq. (7) was given by the core-hole potential from an atomic calculation. For valence calculations we use the potential from a spherical charge centered at $\mathbf{x}$. The volume of the spherical charge is set by the discretization volume of the real-space grid,

$V_x = \Omega / N_x$, the unit cell volume Ω divided by the number of grid points N_x. The screening calculation can be carried out with or without augmentation.

As above, we enforce the symmetry in interchanging $\mathbf{x}$ and $\mathbf{x}'$,

$$W(\mathbf{x}, \mathbf{x}') = \begin{cases} \frac{1}{2}[W^{[\mathbf{x}]}(\mathbf{x}') + W^{[\mathbf{x}']}(\mathbf{x})], & \text{if } |\mathbf{x} - \mathbf{x}'| \leqslant r_m, \\ W_{\mathrm{HLL}}(\mathbf{x}, \mathbf{x}'), & \text{otherwise} \end{cases},$$
$$(14)$$

As for the core-level, the RPA screening is calculated only within a finite space and parametrized by the shell radius R_S and dielectric constant ϵ_∞. At large distances the HLL approximation is still used, governed by the parameter r_m. Much like the shell radius, this constitutes a *controlled* approximation.

Like the HLL model, care must be taken when evaluating the real-space W in the limit of $\mathbf{x} \to \mathbf{x}'$. We numerically integrate the $l = 0$ component of the calculated screened potential over the discretization volume

$$W^{[\mathbf{x}]}(\mathbf{x}) = \frac{3}{R_x^3} \int_0^{R_x} W^{[\mathbf{x}]}(r) r^2 dr, \qquad (15)$$

where $R_x = [3V_x / 4\pi]^{1/3}$.

The system-size scaling behavior is the same for the valence screening as it was for the core-level case—the number of grid points $\mathbf{x}$ scales linearly with volume the same as the expected number of atomic sites. (The scaling is discussed in Sec. VI and illustrated in Sec. VI G.) There are two major differences between calculating the screening for the valence and core cases with negative and positive impacts on runtime. First, the number of real-space grid points is much larger than the number of atoms. For example, in a unit cell of LiF an 10^3 x-point mesh is necessary, resulting in 1000 screening sites instead of the 1 needed for the fluorine x-ray absorption calculation. The dramatic increase in the number of sites is offset by the use of coarser real-space grids in the calculation of χ. The perturbing potential for the core case is the core-hole potential, and, like the core-hole density, it is strongly localized. In the valence case, the perturbing potential is taken to be a uniform ball of charge whose volume is set by the discretization volume V_x defined above. Therefore, the valence screening calculations converge with a coarser radial mesh than is needed for the core-level screening.

B. Effect of model screening and augmentation on optical calculations

To show the effects of using the local RPA, with and without augmentation, instead of the HLL screening on valence calculations we consider calculations of the imaginary part of the dielectric function.

First, we look at bulk silicon, a standard for valence electronic structure calculations providing a testbed for early DFT, GW, and valence BSE calculations [47]. As for the lithium halides we use the experimental lattice constant of 0.543 nm [48] and the PseudoDojo pseudopotential for silicon. Additional input parameters are included in Appendix A. A scissor correction was used to set the DFT band gap to be 1.11 eV [49]. Unsurprisingly, we see in Fig. 10 that in the case of silicon the HLL model performs very well. The inclusion of the RPA screening at short range has little effect on the
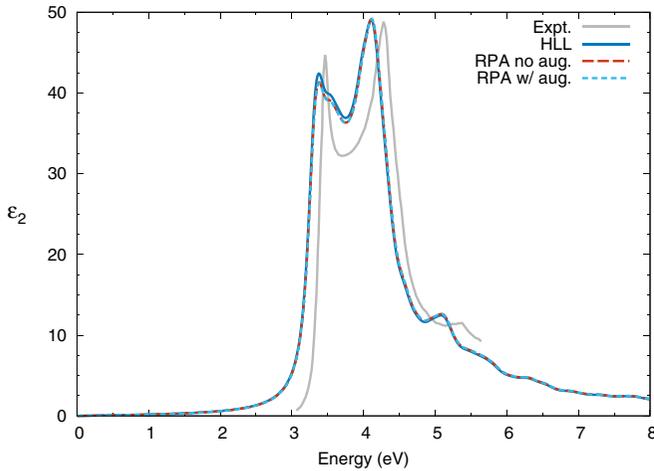
FIG. 10. The calculated imaginary part of the dielectric function for silicon as a function of energy compared with experiment taken from Ref. [45]. The spectra are generated using two different approximations to the screened Coulomb interaction W.

spectrum of bulk silicon, and augmentation has no visible effect.

Next we return to LiF (also well studied previously within the BSE [15,50,51]). The simulation details are similar to the x-ray case, but we have included a scissor correction to set the DFT band gap to 14.2 eV [52]. We first compare our calculations BSE results calculated using ABINIT v. 8.10.2 [53] and yambo v. 4.5.0 [54] codes in Fig. 11(a). A small, 0.09 eV shift in the position of the first exciton between the OCEAN and yambo indicates slightly weaker screening within OCEAN. The ABINIT results were obtained with a scissor-corrected band gap of only 13.3 eV, chosen to align the exciton position with the yambo calculation. The origin of the 1-eV shift in the ABINIT spectrum is not known. The three calculations are largely in agreement with each other in shape and intensity as well as previously published calculations, accounting for the choice of scissor correction.

In contrast to bulk silicon, LiF is strongly ionic with the valence orbitals primarily localized on the fluorine site, making it a more difficult system for the model dielectric response. In Fig. 11(b), we see discrepancies between the HLL-only and RPA results. Both the main exciton near 12 eV and the higher energy peak near 21 eV are red-shifted in the HLL as compared to the RPA calculations. Using the HLL model, the main exciton is at 11.8 eV, while the RPA places it at 12.0 eV. This shift indicates that the HLL model screening is weaker than the RPA calculation for these two peaks. This is consistent with a larger GW band-gap correction that often occurs with the HLL approximation is used as compared to when the RPA is used [55]. As for silicon, the differences in the spectra with and without augmentation are very minor and not distinguishable in the plot.

VI. IMPLEMENTATION WITHIN OCEAN AND $N^2 \log N$ SCALING

Having shown the utility of the local, real-space screening in BSE calculations of both core-level and valence-level
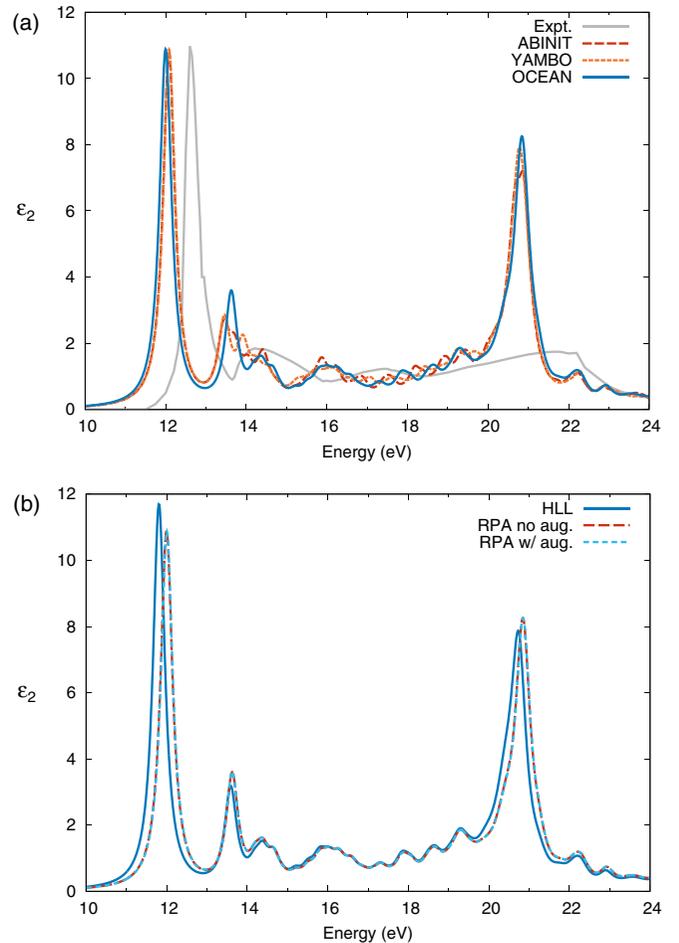

FIG. 11. (a) The imaginary part of the dielectric function for LiF as a function of energy calculated with several different BSE codes and compared to experiment taken from Ref. [46]. The OCEAN calculation was carried out using RPA screening and no augmentation to more closely match the methodology of the other two codes. (b) The LiF dielectric function calculated with OCEAN using different approximations to the screened Coulomb interaction W. The effect of augmentation is minimal, and the two RPA results are nearly identical.

excitations, we now will provide an overview of the implementation and demonstrate the favorable $N^2 \log N$ scaling of our RPA calculations with systems size. As previously stated, our goal is to calculate the screened Coulomb interaction, starting from the irreducible polarizability within the RPA [Eq. (5)]. There are a number of costs and bottlenecks associated with this calculation. To review, the screened Coulomb interaction W is directly calculable from the reducible polarizability χ. The calculation of χ involves matrix products and matrix inversions of the Coulomb operator and the irreducible polarizability χ_0. Within the RPA, χ_0 follows from the one-electron Green's function g, which itself can be written from the the electron orbitals.

In this section, we will explicitly outline the method used in version 3 of OCEAN [15,16] in subsections A through F for each step in calculating the RPA response. We note how data and calculations are distributed for parallel computation, and

we use bars to indicate when a process has only a subset of the total indices, e.g., bands $\bar{b}$ or k-points $\bar{k}$. In subsection G we examine the scaling behavior with system size and parallel performance.

A. Evaluating wave functions and local basis

The initial step is determining the electron wave functions and the basis, from which we can generate the Green's functions. DFT is used to simulate the electronic Hamiltonian. The system is taken to be periodic such that the electron orbitals can be denoted by their band b, crystal momentum $\mathbf{k}$, and spin σ,

$$H^{\text{DFT}}\psi_{b\mathbf{k}\sigma} = \psi_{b\mathbf{k}\sigma}\varepsilon_{b\mathbf{k}\sigma}, \tag{16}$$

where each wave function $\psi_{b\mathbf{k}\sigma}$ has energy $\varepsilon_{b\mathbf{k}\sigma}$. The Green's function for energy E can be written in the spectral representation

$$g_\sigma(\mathbf{r}, \mathbf{r}', E) = \sum_{b=1}^{\infty} \int_{\text{BZ}} \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{\psi_{b\mathbf{k}\sigma}^\dagger(\mathbf{r})\psi_{b\mathbf{k}\sigma}(\mathbf{r}')}{E - \varepsilon_{b\mathbf{k}\sigma}}. \tag{17}$$

The integral runs over the Brillouin zone.

To construct g we must define the real-space basis. As mentioned in Sec. II C, we calculated the RPA response only for the local potential v_1. We employ a real-space basis within a sphere S with a radius r_S and centered on a point τ. For screening the core hole, τ is the atomic site, while for valence calculations τ is one of the grid points in Eq. (14). The irreducible polarizability $\chi_0^{(\tau)}$ is then an $N_r \times N_r$ size matrix, as are the Green's functions $g^{(\tau)}(E)$. This real-space basis is independent of the size of the system's unit cell, and is discussed further in Appendix D 1 a.

In practice the sum over bands is truncated, and the integral is replaced with a sum over regularly spaced points in k-space. Our approach requires only a few k-points, often between 1 and 8, and we address this later in Appendix D 2 on errors and convergence. The sum over bands, however, is significant. Typically, convergence in the screening is reached when the Green's function is constructed with states up to around 100 eV above the Fermi level. This requires on the order of 30 to 50 bands per atom, and the number of bands required scales linearly with system size. Unfortunately, aspects of the generation of the one-electron states from DFT scale with the square of the number of bands.

B. Projecting the wave functions

These DFT eigenstates ψ are generated using an external plane-wave DFT program and are saved to file as Bloch states u,

$$\psi_{b\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{b\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}\sum_{\mathbf{G}} C_{b\mathbf{k}}(\mathbf{G})e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{18}$$

which are defined in terms of complex-valued coefficients C of plane waves $\mathbf{G}$. The spin index σ will be dropped. Only the set of coefficients C, not the various phases, are written to file. We distribute the work for the conversion by band and k-point—not by plane-wave coefficient.

To project the wave functions onto our spherical grid, one option would be to follow the method given by Eq. (18)

directly, i.e., Fourier interpolation. We first create a matrix of the phases, as these will be common across all of the bands at a specific k-point.

$$P_\mathbf{k}(\mathbf{r}, \mathbf{G}) = e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{19}$$

$$\psi_\mathbf{k}(\mathbf{r}, \bar{b}) = \sum_{\mathbf{G}} P_\mathbf{k}(\mathbf{r}, \mathbf{G})C_\mathbf{k}(\mathbf{G}, \bar{b}), \tag{20}$$

where the bar indicates that we are processing only a subset of the total number of bands. The phase matrix requires $\mathcal{O}[N]$ operations from the plane waves, regardless of the number of processors included. Projecting the wave functions is $\mathcal{O}[N^2]$ from the plane waves and bands, but the bands are distributed by processor. The summation over k-points is not counted in the estimation of computational cost because it decreases with volume and is usually 8 or 1. For a system with more than one site of interest, e.g., a disordered, liquid, or amorphous system, the number of sites, and therefore the number of local real-space grids, increases with volume as well. This means that the actual costs increase to N^2 and N^3.

To avoid N^3 scaling, we instead use a fast Fourier transform (FFT), followed by interpolation, and completed by applying the complex phase:

$$u_{\bar{b}\mathbf{k}}(\mathbf{x}) = \mathcal{FFT}[u_{\bar{b}\mathbf{k}}(\mathbf{G})]$$
$$u_{\bar{b}\mathbf{k}}(\mathbf{r}) = \text{Interp.}\,[u_{\bar{b}\mathbf{k}}(\mathbf{x})]$$
$$\psi_{\bar{b}\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\bar{b}\mathbf{k}}(\mathbf{r}). \tag{21}$$

The real-space grid $\mathbf{x}$ is defined as the Fourier transform dual of $\mathbf{G}$. We use third-degree Lagrange polynomial interpolation to generate the wave functions on our desired points $\mathbf{r}$ from the FFT grid $\mathbf{x}$, aided by oversampling the FFT. The interpolants are cached to allow reuse within and between sites. The costs, including a factor of N sites, are $\mathcal{O}[N^2 \log N]$, $\mathcal{O}[N^2]$, and $\mathcal{O}[N^2]$, respectively. All three steps are independent over bands, k-points, and spins, providing good scaling with the number of processors.

To determine the break-even point between these two methods we must be more specific with the actual costs of each step. The N^3 term from method 1 is $N_\mathbf{G}N_bN_\mathbf{r}N_i$, where i are the atomic sites. The $N^2 \log N$ term from method 2 is $A_F N_\mathbf{G}N_b \log N_\mathbf{G}$, where A_F is the FFT prefactor. Therefore, method 2 is preferable if $A_F \log N_\mathbf{G} < N_\mathbf{r}N_i$. Under the assumption that the logarithm of even a large number is about 10, method 2 is likely preferable, even for single-site calculations [56]. More sophisticated methods for Fourier interpolation onto *irregular* grids have been proposed in literature, e.g., Ref. [57] and references therein, but are not explored here.

C. Augmentation

The next step is augmenting the wave functions to recreate the all-electron character,

$$\begin{aligned}\psi_{\bar{b}\bar{\mathbf{k}}}^{\text{ae}}(\mathbf{r}) = &\, \psi_{\bar{b}\bar{\mathbf{k}}}^{\text{ps}}(\mathbf{r}) \\ &+ \sum_{v,l,m} Y_{lm}(\hat{r})[\phi_{vl}^{\text{ae}}(\mathbf{r}) - \phi_{vl}^{\text{ps}}(\mathbf{r})]\langle\phi_{vl}^{\text{ps}}Y_{lm}|\psi_{\bar{b}\bar{\mathbf{k}}}^{\text{ps}}\rangle,\end{aligned}$$

$$\tag{22}$$

where ϕ are the OPFs and Y_{lm} are spherical harmonics. The local basis is substantially coarser than that used in the construction of the OPFs. Therefore we enforce unitarity by constructing the overlap matrix A,

$$A_{vv';l} = \int_0^{r_a} dr\, r^2 \phi_{vl}^{ps}(r) \phi_{v'l}^{ps}(r), \quad (23)$$

where any deviation from the identity matrix is due to errors from using a coarser grid. The augmentation of Eq. (22) is modified,

$$\psi_{b\mathbf{k}}^{ae}(\mathbf{r}) = \psi_{b\mathbf{k}}^{ps}(\mathbf{r}) + \sum_{lm} \sum_{v,v'} Y_{lm}(\hat{r}) \big[\phi_{vl}^{ae}(\mathbf{r}) - \phi_{vl}^{ps}(\mathbf{r}) \big]$$
$$\times A_{vv';l}^{-1} \langle \phi_{v'l}^{ps} Y_{lm} | \psi_{b\mathbf{k}}^{ps} \rangle, \quad (24)$$

preserving unitarity.

Each process stores a copy of the OPFs and carries out the augmentation for its subset of bands and k-points. The scaling of this section goes as $\mathcal{O}[N^2]$ with a factors of N_i sites and N_b bands.

D. Building g and χ_0

The RPA polarizability from Eq. (5) can be transformed from the two-time form to a convolution over energy, which can be carried out along the imaginary axis,

$$\chi_0(\mathbf{r}, \mathbf{r}', \omega = 0)$$
$$= -i \sum_\sigma \int_{-\infty}^\infty \frac{dE}{2\pi} g_\sigma(\mathbf{r}, \mathbf{r}', E) g_\sigma(\mathbf{r}', \mathbf{r}, E)$$
$$= \sum_\sigma \int_{-\infty}^\infty \frac{dt}{2\pi} g_\sigma(\mathbf{r}, \mathbf{r}', \mu + it) g_\sigma(\mathbf{r}', \mathbf{r}, \mu + it), \quad (25)$$

where μ is chosen to be in the middle of the gap for insulators or at the Fermi level in metals to avoid poles in g. [With minimal approximation, a small energy Δ can be added in quadrature in metals to the difference between μ and the Kohn-Sham eigenvalue $\varepsilon_{b\mathbf{k}}$ according to $(\mu - \varepsilon_{b\mathbf{k}}) \to \pm\sqrt{(\mu - \varepsilon_{b\mathbf{k}})^2 + \Delta^2}$.] This same approach can be used for calculating the dynamic polarizability, $\omega \neq 0$. However, additional care is needed to avoid the poles in Green's function $g(\omega + \mu + it)$ as $t \to 0$. The integral is replaced by a sum over an energy grid as outlined in Appendix D 1 b.

In principle, partial Green's functions could be constructed using the band and k-point distribution from the previous step. However, it is more efficient to redistribute the wave functions into blocks by $\mathbf{r}$ and site. The processors are split into groups such that each group works on its own site or set of sites. Within each group, the processors divide up the $\mathbf{r}$-points. This means that the wave functions are now distributed as $\psi_{b\mathbf{k}\sigma}(\bar{\mathbf{r}}; \bar{\tau})$,

$$g_\sigma^{(\tau)}(\bar{\mathbf{r}}, \bar{\mathbf{r}}', \mu + it) = N_k^{-1} \sum_{\mathbf{k}} \sum_{b=1}^{N_b} \frac{\psi_{b\mathbf{k}\sigma}^\dagger(\bar{\mathbf{r}}) \psi_{b\mathbf{k}\sigma}(\bar{\mathbf{r}}')}{\mu + it - \varepsilon_{b\mathbf{k}\sigma}}, \quad (26)$$

where $\mathbf{r}$ implicitly includes only the points for site τ. If background communications are enabled, then the majority of this data transfer takes place while the conversion process is ongoing. Within a group of processors cooperatively working on a site $g^{(\tau)}$, the wave functions are shared. The scaling of

this section goes as $\mathcal{O}[N^2]$ with a factors of N_i sites and N_b bands.

An important consideration in efficiently calculating the Green's function is that it involves an outer product of the wave functions. For each band and k-point, N_r inputs are turned into N_r^2 outputs, with $2N_r^2$ operations. In a typical, small calculation the real-space grid has 1600 points per site, which means that at each frequency the Green's function is just under 40 MB in size, making it too large to fit in the local cache of a typical modern processor. A naïve implementation would be limited by memory bandwidth. Instead, the wave functions are broken up, and the Green's function is calculated by tiles in real space.

E. Construction of χ and W

In the previous step, we calculated the irreducible polarizability χ^0. The reducible polarizability is given by

$$\chi = [1 - \chi^0 v]^{-1} \chi^0, \quad (27)$$

where v is the Coulomb potential operator. We do this by projecting into a spherical basis

$$\chi^0(\mathbf{r}, \mathbf{r}') = \sum_{lm} \sum_{l'm'} \chi_{lm;l'm'}^0(r, r') Y_{lm}(\hat{r}) Y_{l'm'}^*(\hat{r}'), \quad (28)$$

$$v(\mathbf{r}, \mathbf{r}') = \sum_{lm} \frac{4\pi}{2l+1} \frac{r_<^l}{r_>^{l+1}} Y_{lm}(\hat{r}) Y_{lm}^*(\hat{r}'), \quad (29)$$

where the Coulomb operator is diagonal in l, m. We can define

$$S_{lm;l'm'}(r, r') = \delta_{l,l'} \delta_{m,m'} \delta(r - r')$$
$$- \int dx\, x^2 \chi_{lm;l'm'}^0(r, x) \frac{4\pi}{2l'+1} \left[\frac{r_<^{l'}}{r_>^{l'+1}} \right]_{x,r'} \quad (30)$$

by taking advantage of the diagonal nature of v in this basis. We therefore have

$$\chi_{lm;l'm'}(r, r') = S_{lm;l''m''}^{-1}(r, r'') \chi_{l''m'';l'm'}^0(r'', r'), \quad (31)$$

where the matrices have dimension $N_r(2N_l + 1)$.

In this basis the induced change in electron density from the short-range part of the core-hole potential $v^{(1)}$ is

$$\rho^{ind}(\mathbf{r}) = \sum_{lm} \rho_{lm}^{ind}(r) Y_{lm}(\hat{r}), \quad (32)$$

$$\rho_{lm}^{ind}(r) = \sum_{l'm'} \int d^3r'\, \chi_{lm;l'm'}(r, r') v^{(1)}(r') Y_{l'm'}(\hat{r}')$$
$$= \int dr'\, r'^2 \chi_{lm;00}(r, r') v^{(1)}(r'). \quad (33)$$

The perturbing (core-hole) potential is taken to be spherical and therefore only the $l' = 0$ part of χ contributes. Giving a final, induced potential

$$v^{ind}(\mathbf{r}) = \sum_{lm} v_{lm}^{ind}(r) Y_{lm}(\hat{r})$$

$$v_{lm}^{ind}(r) = \int dr'\, r'^2 \rho_{lm}^{ind}(r') \frac{4\pi}{2l+1} \frac{r_<^l}{r_>^{l+1}}. \quad (34)$$

TABLE I. Timing in seconds of selected LiF runs, denoted by the number of fluorine atoms and the k-point sampling (k-pts) running on 128 processors (see text). The DFT timing includes only the "electron" time for the non-self-consistent run to generate the wave functions for the screening. The two divisions of the screening calculation, labeled ψ and $g\,\&\,\chi$, encompass the totality of the runtime for the screening. The timing for ψ includes the time to read in the wave functions, project and augment them, and redistribute them. The timing for $g\,\&\,\chi$ includes the time to construct the Green's functions, evaluate the polarizability, and screen the core-hole potential. The root-mean-square deviations were determined over eight repeated runs and are included in parentheses.

N_F	Vol. (a.u.^{3})	k-pts	DFT (s)	ψ (s)	$g\,\&\,\chi$ (s)
384	42002	Γ	27421.6	333.6 (8.4)	1474 (22)
288	31502	Γ	11556.1	208.1 (4.5)	891.0 (14.6)
216	23626	Γ	6063.4	106.4 (3.0)	445.7 (5.2)
125	13673	Γ	1350.2	36.1 (0.1)	131.0 (1.5)
64	7000	Γ	189.6	13.7 (0.1)	36.5 (1.1)
27	2953	Γ	31.2	5.36 (0.02)	13.0 (0.4)
125	13673	1	2791.2	92.0 (1.0)	166.1 (3.6)
64	7000	1	609.5	79.9 (0.2)	57.0 (5.1)
27	2953	1	62.5	30.3 (0.1)	15.6 (0.5)
8	875	1	6.0	9.36 (0.41)	1.28 (0.04)
64	7000	8	4585.4	175.5 (5.2)	425.4 (72.9)
27	2953	8	408.6	45.8 (1.0)	128.0 (4.7)
8	875	8	17.9	17.5 (0.3)	9.86 (0.03)
1	110	8	0.2	4.39 (0.11)	0.46 (0.04)

By default, only the $l = 0$ part of the response is calculated, according to

$$\bar{\chi}(r, r') = S^{-1}_{00;00}(r, r'')\chi^0_{00;00}(r'', r')$$

$$\bar{\rho}^{\text{ind}}(r) = \int dr' r'^2 \bar{\chi}(r, r')v^{(1)}(r'). \qquad (35)$$

The resulting induced potential is approximately the same as the $l = 0$ component of the full induced density $\bar{\rho} \approx \rho_{00}$. For the core-hole potential, the strong localization means that $l = 0$ component of the induced potential is dominant, and this approximation is reasonable. For valence calculations, including $l \leqslant 2$ was found to be sufficient. Because the dimension of χ is independent of system size, this section scales only with the number of sites N_i, linearly with system size or $\mathcal{O}[N]$.

F. Γ-point

For large cells, only a single k-point is required, and the electron orbitals can be calculated at the Γ-point. For systems with time-reversal symmetry the Bloch functions can be treated as real (instead of complex). This results a reduction of the required storage by half and substantial time savings in the DFT stage. A smaller reduction in runtime is also realized in the screening calculation as shown below in Table I.

G. Timing and scaling

The calculation of the screening as outlined here is dominated by three steps: calculating the wave functions, pro-
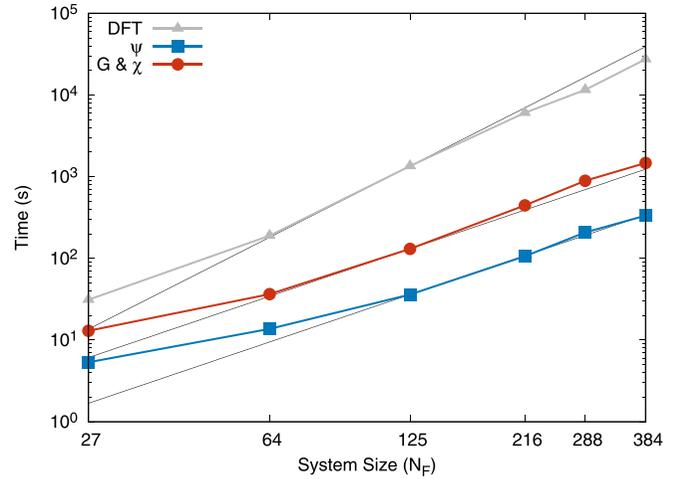

FIG. 12. The runtime for the DFT and screening as a function of the system size. The screening is divided into the wave function projection ψ (blue, squares) and construction of the Green's function and polarizability $G\,\&\,\chi$ (red, circles). Guidelines are shown for αN^2, normalized to $N_F = 125$. For the DFT calculation (gray, triangles) the guideline is N^3. The plotted data are included in the top section of Table I.

jecting them onto the radial grid, and constructing the Green's function and χ^0. The first, calculating the electron orbitals using DFT, is carried out using the QUANTUM ESPRESSO code [22,23]. We report the timing of the DFT step for completeness, however, we are focused on the two steps that are specific to the screening calculation.

To investigate the timing and scaling of screening calculations within OCEAN we use LiF (physical details and convergence parameters are given in Appendix A). There are two classes of scaling that we are interested in. First there is system scaling, by which we mean the increase in runtime with an increase in the system size. This highlights the inherent simulation size limits of our approach. We will also consider strong scaling, the change in execution time due to changing the number of processors. We have implemented two levels of parallelism for the screening calculations: internode MPI and shared memory OPENMP. The testbed for these calculations is a small cluster with 12 nodes. Each node has a dual-socket with 8 processors per CPU (16 per node) [58]. Each timing run was repeated 8 times, and the average value is reported (DFT calculations were run only once).

For the tests, we consider various supercells of LiF, from the unit cell $N_F = 1$, to an $6 \times 8 \times 8$ supercell $N_F = 384$, covering cell volumes from 110 a.u.^{3} to 42000 a.u.^{3} (6.2 nm^3). For these runs, 32 bands per unit cell were included (12 288 for $N_F = 384$). For each supercell the screened Coulomb potential is calculated for all the fluorine sites. Each local real-space grid had 2624 points, and the Green's functions were evaluated at 16 imaginary frequencies.

First we look at the scaling with system size or weak scaling. In Fig. 12 we show the runtime of the projection (ψ) and construction of the Green's function and polarizability ($g\,\&\,\chi$) steps as a function of super cell size. The straight lines on the log-log plot are αN^2, where α is set by the timing of $N_F = 125$. In this set of runs only Γ-point sampling of the

Brillouin zone was used. The $\approx \mathcal{O}[N^2]$ growth in calculating the screening is evident by the linear plots, though overhead or inefficiencies dominate the timing of the smallest run. Additionally, for large systems the "g & χ" diverges slightly from the expected N^2 behavior which may be indicative of poor cache reuse or other bottlenecks for large system sizes. We also include the timing of the DFT portion which follows αN^3. One hundred twenty-eight processors across eight nodes were used for each run, using 128 MPI processes.

The timing information for the range of systems from $N_F = 1$ to $N_F = 384$ is shown in Table I. To give a full picture of the scaling, three different settings for the k-point sampling were used: finite sampling on a 2^3 k-point mesh, single k-point sampling, and Γ-point sampling. Above $N_F = 27$, a single k-point is sufficient to sample the Brillouin zone and gives the same results as the 2^3 sampling (but more quickly). The purpose of timing single k-point runs in addition to the Γ-point runs is to distinguish the changes due to the reduction in k-point sampling from 8 to 1 from the changes in moving between complex and real Bloch functions. The single k-point is taken at ($1/8$, $2/8$, $3/8$).

Next, we present the change in runtime with changing processor number or strong scaling, Fig. 13. Ideally, doubling the number of processors used in a calculation will halve the runtime. Longer than expected runtimes may result from serial sections of the code or communication overhead. Shorter than expected times may result from better data caching due to each processor working on a smaller data set. Here we plot the data as the efficiency E as a function of the number of processors N,

$$E(N) = \frac{N_0}{N} \frac{t(N_0)}{t(N)} \times 100\%, \tag{36}$$

where the efficiency is normalized to the runtime with N_0 processors. Ideal scaling is given by an efficiency of 100%. The efficiency is the measure of merit for planning high-throughput calculations. In high-throughput calculations the available hardware resources can be divided between many different calculations, and the runtime of any single calculation should be balanced against the runtime of the complete dataset.

In Fig. 13(a) we show that for a moderately sized system $N_F = 64$, there is a drop-off in efficiency above 16 processors. In part this is a reflection of the structure of our computer cluster, where each addition of 16 processors increases the number of nodes in the calculation by 1. While the efficiency is quite poor running on 160 processors, the runtime is also very brief. The average time is 41.1 s compared to an idea time of 27.3 s. The larger systems show better scaling.

Last, we examine the OMP parallelism by repeating the Γ-centered calculations this time including thread-level parallelism. The total number of processors is held fixed at 128, but divided between MPI and OMP with 1, 2, 4, and 8 OMP threads per MPI process. The results are shown in Table II and for three of system sizes in Fig. 13(b). We find relatively uniform performance across the first three processor arrangements, but a drop-off in performance using 8 OMP threads. This drop-off indicates an opportunity for further code refinement to better support higher levels of OPENMP parallelism. In Table II it can
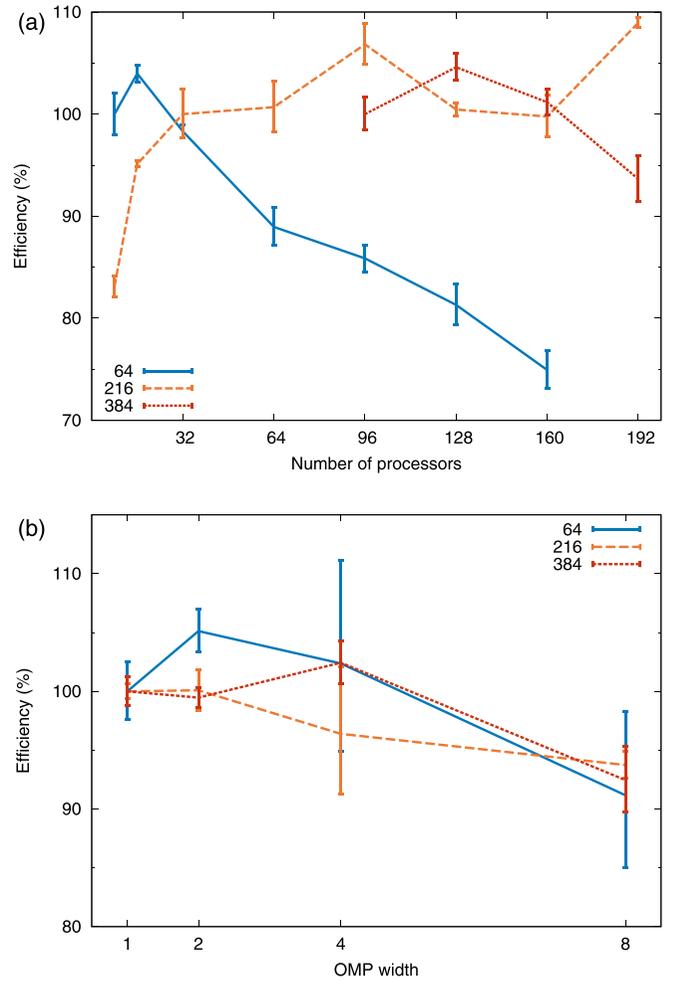

FIG. 13. (a) The strong scaling behavior of the complete screening calculation for three different systems sizes, $N_F = 64$ (solid, blue) normalized for 8 processors, $N_F = 216$ (dashed, orange) normalized for 32 processors, and $N_F = 384$ (dotted, red) normalized at 96 processors. The error bars reflect the root-mean square deviation determined by averaging over 8 runs. We see a falloff in efficiency of the smaller system, but for $N_F = 216$ and $N_F = 384$ we see good strong scaling up to 192 processors. (b) The efficiency of the OMP parallelization, normalized with one OMP thread per MPI process. The number of processors is kept fixed at 128 and divided between MPI and OMP.

be seen that this inefficiency is primarily the result of poor scaling of the "ψ" section.

VII. DISCUSSION AND FUTURE DIRECTIONS

We have presented a local, real-space method for calculating the RPA polarizability of condensed systems. The method scales well with system size N, $\mathcal{O}[N^2 \log N]$. While the method only provides the full RPA response within a restricted real-space range, it is coupled with a model dielectric function to provide the full response. This approximation is controlled through a radial cutoff R_S, and the contribution of the model goes smoothly to zero as $R_S \to \infty$. This method is implemented within the OCEAN code where the screened Coulomb operator W is used as part of the BSE Hamiltonian for

TABLE II. Timing in seconds of LiF runs with Γ-point sampling. A total of 128 processors are used for each run divided between MPI tasks and OPENMP threads. The two timing sections are the same as in Table I.

N_F	MPI	OMP	ψ (s)	g & χ (s)
384	128	1	333.6 (8.4)	1474 (22)
	64	2	330.9 (11.6)	1488 (15)
	32	4	336.4 (2.6)	1430 (31)
	16	8	504.5 (19.3)	1452 (49)
288	128	1	208.1 (4.5)	891.0 (14.6)
	64	2	204.2 (4.9)	898.1 (33.0)
	32	4	210.4 (2.1)	767.4 (5.7)
	16	8	315.0 (12.6)	775.8 (6.0)
216	128	1	106.4 (3.0)	445.7 (5.2)
	64	2	105.8 (2.1)	447.1 (9.3)
	32	4	108.7 (5.4)	465.9 (28.4)
	16	8	167.3 (2.1)	423.6 (7.4)
125	128	1	36.1 (0.1)	131.0 (1.5)
	64	2	35.4 (1.0)	151.7 (2.9)
	32	4	34.3 (0.2)	148.0 (7.2)
	16	8	53.6 (0.7)	157.4 (13.3)
64	128	1	13.7 (0.1)	36.5 (1.1)
	64	2	11.6 (0.1)	37.4 (0.8)
	32	4	12.4 (1.7)	38.6 (4.1)
	16	8	16.3 (0.1)	40.6 (4.4)
27	128	1	5.36 (0.02)	13.0 (0.4)
	64	2	2.88 (0.01)	8.4 (0.5)
	32	4	2.69 (0.36)	10.4 (1.7)
	16	8	3.35 (0.02)	12.2 (0.3)

calculating both core-level (near-edge x-ray) and valence-level (UV/vis) spectroscopy.

In regions near an atom, we have shown that the pseudopotential approximation results in an incorrect RPA polarizability. In our screening calculations we correct for this by augmenting the electron orbitals from pseudopotential-based calculations to restore the all-electron character. This effect is noticeable in near-edge x-ray spectra, where changes in exciton strength and position due to deficiencies in the screened core-hole potential are similar to changes due to thermal disorder. Above the x-ray edge the differences between spectra calculated with and without augmented orbitals fade with increasing photoelectron energy as the photoelectron becomes more and more delocalized. In the case of valence-level UV/vis spectroscopy we found that augmentation is not necessary.

We conclude with a few remarks on improvements and future extensions. In particular, the relative ease of the local, real-space method may present an opportunity for developing and testing new model dielectric functions and easily benchmarking them against RPA or time-dependent DFT quality calculations. In the remainder of this section we first detail some enhancements to the current method. Next, we discuss computations other than particle-hole spectroscopy that could benefit from our local, real-space polarizability. Finally, we show how the real-space method is amenable to higher-order calculations beyond RPA.

A. Refinements

In the current implementation there is no re-use of the Green's functions between sites. The use of site-centered radial grids makes it unlikely that a given $(\mathbf{r}, \mathbf{r}')$ pair of one site will exist in the grid of another. However, it reasonable to expect that many point pairs will be nearly shared, e.g., for sites α and β that $\mathbf{r}_\alpha \approx \mathbf{r}_\beta$ and $\mathbf{r}'_\alpha \approx \mathbf{r}'_\beta$. In the future, the construction of the grids can be relaxed to maximize the overlaps, decreasing the computational cost of generating the Green's functions. This would be especially helpful in the case of valence calculations where the site density is high.

Future improvements to the scalability with system size must focus on generating the electron wave functions. For medium to large system sizes, most of the time in calculating the screening is spent in the DFT (see Sec. VI G). This is exacerbated by the need for unoccupied states in the calculating the Green's function. Several methods have been proposed to reduce the number of unoccupied states.

One option is to directly replace part of the sum over unoccupied states. The effective-energy techniques replace the energy denominator in the sum over unoccupied states [59–61]. The completeness of the Bloch functions then allows the sum over unoccupied bands to be replaced with the identity minus a sum over occupied bands. However, these approaches differ in two main ways from ours, not including our local approximation. First, the energy convolution is carried out analytically, and the RPA polarizability is constructed via sums over states. In our approach the Green's functions are built via a sum over states and the convolution is carried out numerically. Second, the effective-energy approaches are formulated in reciprocal space, which has the advantage of a straightforward approximation for the effective energy. An easier approach might be to approximate the neglected high-energy bands as plane waves [62–65].

Alternatively, the induced density response and therefore the screened potential can be more directly calculated using the linear-response Sternheimer equation approach [5,66] or eigenvalue decomposition of the polarizability matrix [6,67]. While these approaches only require the occupied orbitals, they maintain an unfavorable N^4 scaling with system size. Better scaling might be achievable by adapting these approaches to determine only the local response.

Last, the model dielectric function used to screen v_2 [Eq. (10)] was designed for bulk systems. In the case of highly anisotropic systems, like a surface or interface, this may result in slower convergence with respect to the shell radius R_S, requiring a larger real-space RPA calculation. This could be addressed by modifying the screening model or alleviated through a judicious choice of model parameters, namely choosing the average density ρ_0 to more accurately reflect the bulk material (see model details in Appendix B).

B. Non-BSE applications

The screened Coulomb interaction W has many uses in calculations of condensed matter systems other than the use presented here of the direct interaction in the the BSE. One such application is in self-energy calculations using the GW method which requires evaluating the self-energy operator

[68],

$$\Sigma(\mathbf{r}, \mathbf{r}', E) = i \int \frac{d\omega}{2\pi} G(\mathbf{r}, \mathbf{r}', E - \omega) W(\mathbf{r}, \mathbf{r}', \omega). \quad (37)$$

The local screening approach outlined here can be used to efficiently generate W with RPA quality for small distances $|\mathbf{r} - \mathbf{r}'|$, just as was shown for valence BSE calculations. The terms in the GW calculations are significantly less localized than core-level excitations, making discrepancies in the short-range part of W less important, and augmentation of the orbitals may not be necessary. However, in transition metals the d orbitals often drive important characteristics of the electronic behavior, forming the top of the valence bands, the bottom of the conduction bands, or both. From atomic calculations, it can be seen that the d orbitals overlap significantly with the semicore orbitals of the same principal quantum number, e.g., the $3d$ with the $3s$ and $3p$. High-accuracy calculations involving localized d orbitals may require accurately correcting the nodal structure of the s and p wave functions in much the same manner as the we have shown for core-level spectroscopy. Several GW studies have pointed out discrepancies from using pseudopotentials [36] or excluding semicore states [69,70]. In the present approach only static screening was implemented. However, the contour integral in Eq. (25) can be modified to calculate $\omega \neq 0$, and the fundamental scaling of frequency-dependent screening remains the same as that of the static case. In addition to standard valence- and conduction-band self-energy calculations, an adaptation of this method could be applicable for determining accurate core-level binding energies [71].

The local, real-space screening might also be useful for phonon calculations. Within the harmonic approximation, the phonons of a system can be fully described by the dynamical matrix. The elements of the dynamical matrix are proportional to the derivative of the force on atom a with respect to changes in the position of a'. This is equivalent to the second derivative of the total energy with respect to the displacement of both. The elements of the dynamical matrix can be calculated using density-functional perturbation theory [72–74]. Alternatively, the dielectric response can be used since the polarizability describes how the electron density will change in response to a change in the potential, in this case the motion of the atomic nuclei. Care would be required as the polarizability gives the density change in response to a local perturbing potential, but standard pseudopotentials include nonlocal terms [75].

C. Beyond RPA screening

The screening calculations in this paper have been carried out only within the RPA approximation. From a many-body perturbation theory perspective, the RPA is the lowest-order diagram for the polarizability. Higher-order approaches treat interactions between the electron and hole lines in the RPA, or, equivalently, add a vertex correction. Unfortunately, additional interaction or vertex terms increase the computational cost and scale worse with system size. Our local, real-space approach is an ideal testbed for investigating higher-order approaches because the increase in scaling complexity applies to the dimension of our local dielectric response function which is small and independent of the system size.

As an example, we have implemented the vertex correction given by the adiabatic local-density approximation (ALDA). As shown by Del Sole, Reining, and Godby [76], if the first-order approximation to the one-electron self-energy is taken to be the local exchange-correlation potential

$$\Sigma(1, 2) = \delta(1, 2)v_{\text{xc}}(1), \quad (38)$$

then the reducible polarizability (and screened interaction W) undergo a relatively simple transformation. Repeating Eq. (6),

$$\chi = (1 - \chi_0 v)^{-1} \chi_0$$
$$\tilde{\chi} = [1 - \chi_0(v + f_{\text{xc}})]^{-1} \chi_0, \quad (39)$$

where f_{xc} is the derivative of the exchange-correlation potential with respect to the density, $f_{\text{xc}}(1, 2) = \delta v_{\text{xc}}(1)/\delta n(2)$, and $\tilde{\chi}$ is the ALDA polarizability. The use of an ALDA kernel has been investigated within the GW approximation [76–78] and for valence BSE calculations of small molecules [79].

Within the ALDA, f_{xc} is a contact interaction, and the expression in Eq. (39) is easily evaluated using the OCEAN code as outlined in Sec. VI E. In the real-space basis f_{xc} is diagonal and can be written

$$f_{\text{xc}}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \frac{d v_{\text{xc}}(n)}{dn}\bigg|_{n=n(\mathbf{r})}, \quad (40)$$

where v_{xc} is the LDA exchange-correlation potential and is evaluated at the density n at position $\mathbf{r}$. The electron density $n(\mathbf{r})$ is taken from the initial DFT calculation used to generate the electron orbitals for the screening. We start with the Perdew-Zunger parametrization for the exchange [80] and Vosko, Wilk and Nusair parametrization of the correlation energy [81] within the local-density approximation fit to the data of Ceperley and Alder [82]. We calculate f_{xc} directly as the second derivative of the exchange-correlation energy with respect to the density using a five-point finite difference using density differences of 0.01 e^- per a.u.^{3} Spin-polarized calculations beyond the RPA are not yet supported, but can be included using this same scheme.

Once again looking at h-BN, we can examine how calculating the polarizability with the ALDA instead of RPA changes the XAS. In general, the ALDA results in a stronger induced potential [shown for LiF in Fig. 16(a)]. This in turn leads to a weaker core-hole potential and correspondingly weaker excitonic effects. In Fig. 14 we show the nitrogen K-edge XAS using both the RPA and ALDA for the screening. The BSE spectrum calculated using the ALDA is substantially different from the RPA result, but only in the near-edge region, within about 15 eV of the onset. The small differences at higher energies would be hidden by broadening if the calculation included the effects of the electron self-energy and vibrational disorder. While electron energy loss spectroscopy (EELS) taken within the dipole limit should probe the same excitations as XAS, it is seen in h-BN that there are large discrepancies between the two [40]. In part, this is due to the different systematic errors such as surface sensitivity and self-absorption effects affecting XAS versus EELS. The ALDA screening appears superior to the RPA when comparing to the EELS data, but the RPA appears superior when comparing to the XAS measurement. A broad survey of materials and careful quantification of experimental uncertainties is necessary to
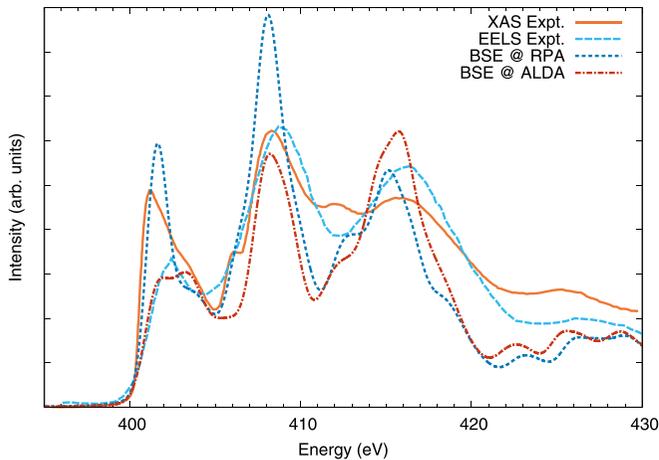
FIG. 14. The N K-edge absorption for *h*-BN calculated using BSE, but two different approximations to calculate the screening. The ALDA approximation screens the core hole much more efficiently [see Fig. 16(a)] resulting in significantly less excitonic binding. Compared to the RPA results, the ALDA has a weaker exciton and a general shift of spectral weight to higher energies. Two experimental spectra are included showing a large variation between XAS [39] and electron energy loss spectroscopy (EELS) [40].

establish the general applicability of one approximation or the other and should be the subject of future work.

While the only vertex correction that has been implemented in OCEAN is the ALDA, an extension to semilocal exchange-correlation kernels is straightforward, requiring only the additional knowledge of the density gradients. Because f_{xc} is formed explicitly in real-space, the formulation of Eq. (39) is also compatible with nonlocal exchange-correlation potentials. This would require construction of f_{xc} as a real-space matrix instead of the diagonal form [Eq. (40)]. However, the response is still localized and in response to a local perturbing potential with the long-range response handled by a model. Therefore, any nonlocal f_{xc} must also be of limited range.

APPENDIX A: INPUT PARAMETERS FOR X-RAY AND OPTICAL CALCULATIONS

We consider the halide K edges of lithium halides, LiF, LiCl, LiBr, and LiI. All crystallize in the same rock salt $Fm\bar{3}m$ structure with lattice constants of 0.4017, 0.5130, 0.5501, and 0.6000 nm, respectively [48]. The plane-wave cut-off energy was set to 100 Ry. (increased to 120 Ry. for the bromine and iodine pseudopotentials), and the density was converged using a 4^3 shifted *k*-point grid. The BSE final states were solved on a 15^3 grid (16^3 for LiI), including 32, 59, 127, or 128 unoccupied bands, respectively, and were downsampled onto a 12^3 real-space mesh (10^3 for LiF). The calculations used the local-density approximation for the density functional [24], and pseudopotentials are taken from PseudoDojo [25] and generated with ONCVPSP [26]. The DFT calculations were carried out using QUANTUM ESPRESSO [22,23]. We used the "high-accuracy" version of the lithium pseudopotential, which includes the Li 1*s* as valence. For bromine and iodine the standard 3*d*4*s*4*p* (4*d*5*s*5*p*) were used, and additional calculations

were carried out with the semicore iodine pseudopotential. Note that no valence-level spin-orbit coupling is considered, which would affect the Br 4*p* or I 5*p* states.

For the screening calculations of the lithium halides, the orbitals for the screening calculation were generated on a 4^3 *k*-point grid, including 72, 150, 197, or 213 bands, for F, Cl, Br, and I respectively, such that energy range from the Fermi level (mid gap) to the highest unoccupied state was approximately 150 eV for each. The augmentation radius of each was set by the pseudopotential of each halide, 1.64, 1.76, 1.97, and 2.02 a.u., respectively, and 1.45 a.u. for the I semicore. For the heavier three the polarizability was calculated within a sphere of radius 8 a.u. on a 160-point uniform radial grid and 64-point angular grid while the neutralizing shell was placed at $R_S = 4$ a.u. For the LiF the polarizability was calculated within a sphere of radius 10 a.u. with the neutralizing shell placed at $R_S = 6$ a.u. The real-space grid was divided into three sections. The inner section used a 34-point Gauss-Legendre quadrature for the radial grid and a 64-point angular grid. From the augmentation radius of 1.64 to 2.96 a.u. a 27-point, uniformly spaced radial grid and 144-point angular grid was used, and the final grid was an 88-point, uniformly spaced radial grid and 256-point angular grid. This grid is excessive for calculations of spectra, but was chosen to accurately show the convergence effects in Fig. 16. For the scaling tests of Sec. VI G a smaller grid was used. A 16-point uniform radial mesh and 36-point angular mesh was used for up to the augmentation radius, a 32-point uniform radial mesh and 64-point angular mesh was used outside it, and the sphere radius was limited to 8 a.u., giving a total of 2624 points.

For hexagonal boron nitride we used the experimentally determined lattice constants of $a = 2.504$ Å and $c = 6.661$ Å [48]. As for the lithium halides, pseudopotentials were taken from the PseudoDojo collection with a plane-wave cutoff of 100 Ry. A $20 \times 20 \times 8$ *k*-point mesh was used for the BSE with 92 unoccupied bands, while the screening was carried out using a $6 \times 6 \times 2$ *k*-point mesh and 300 bands. The EXCITING calculation was carried out using a $10 \times 10 \times 6$ *k*-point mesh and 20 unoccupied bands. A sphere of radius 12 a.u. was used for the polarizability, divided into 5 sections with cutoffs of 1.39, 3, 4, 5, 9, and 12 a.u.; 17-, 10-, 5-, 14-, and 6-point radial sampling; and angular meshes of 36, 100, 256, 625, 576, and 100. These larger grids were to ensure that the small variations in the potential (Fig. 9) were not due to numerical deficiencies and to accommodate the large shell radius $R_S = 7$ a.u.

For the valence calculations, the LiF used a 10^3 real-space grid for the BSE final states, requiring RPA screening calculations on that grid. The polarizability was calculated within a sphere of radius 10 a.u. and the neutralizing shell was placed at $R_S = 4$ a.u. For the LiF valence calculations BSE final states were calculated on a 8^3 *k*-point mesh with 6 conduction bands and 5 valence bands. This lower *k*-point mesh was chosen to facilitate the more computationally expensive comparison calculations. The RPA screening for the valence used a 3^3 *k*-point mesh and 72 bands. Silicon crystallizes in a $Fd\bar{3}m$ structure with experimental lattice constant 0.543 nm [48]. The PseudoDojo pseudopotential for silicon and a plane-wave cutoff of 100 Ry were used. The BSE final states were calculated on a 16^3 *k*-point grid, including 8 conduction and 4 valence bands, and a real-space grid of 8^3. For the RPA

screening 200 bands on a 2^3 k-point mesh were used, and the polarizability was calculated with a sphere of radius 8 a.u. with the neutralizing shell placed at $R_S = 3.5$ a.u.

APPENDIX B: MODEL POLARIZATION

Here we reproduce the model screening of a spherical shell of charge by the model dielectric function introduced in Ref. [14] as used in Ref. [11]:

$$\chi_M(\mathbf{r}, \mathbf{r}') = (2\rho_0)^{-1}$$
$$\times \nabla \cdot \nabla'[[\rho(\mathbf{r}) + \rho(\mathbf{r}')]B(|\mathbf{r} - \mathbf{r}'|)], \quad \text{(B1)}$$

where ρ is the local electron density and ρ_0 is the average electron density. The real-space model B is transform of the Levine-Louie dielectric model $B(q)$.

$$B(|\mathbf{r} - \mathbf{r}'|) = \int \frac{d^3q}{(2\pi)^3} B(q) \exp[i\,\mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')], \quad \text{(B2)}$$

$$B(q) = \frac{1}{4\pi}\left[\frac{1}{\epsilon_{LL}(q; \rho_0, \epsilon_\infty)} - 1\right]. \quad \text{(B3)}$$

The original formulation of the Levine-Louie model in Ref. [13] requires something akin to an average band gap, but this can be reformulated using the long-range dielectric constant ϵ_∞.

$$\epsilon_{LL} = 1 + \frac{2}{\pi q_F}\left\{\frac{1}{Q^2} - \frac{\lambda}{2Q^3}\left(\tan^{-1}\left[\frac{2Q + Q^2}{\lambda}\right]\right.\right.$$
$$+ \tan^{-1}\left[\frac{2Q - Q^2}{\lambda}\right]\right) + \left(\frac{\lambda^2}{8Q^5} + \frac{1}{2Q^3} - \frac{1}{8Q}\right)$$
$$\left.\times \ln\left[\frac{\lambda^2 + (2Q + Q^2)^2}{\lambda^2 + (2Q - Q^2)^2}\right]\right\}, \quad \text{(B4)}$$

where $Q = q/q_F$, $\lambda^2 = (\epsilon_\infty - 1)^{-1}\omega_p^2\omega_F^{-2}$, ω_p is the plasmon frequency, and ω_F and q_F are the Fermi frequency and wave vector of a noninteracting electron gas of density ρ_0.

APPENDIX C: CONSTRUCTION OF OPFS

The construction of the projectors is as follows. For each angular momentum l, self-consistent solutions to Eq. (11) are determined for both the all-electron and pseudopotential systems, e.g., for an isolated atom in either the ground state or a positive ion such that the valence electrons are all bound. For this purpose, the desired energy window for the projectors is selected by choice. For each l, the energy of the most-bound valence state is found. In some cases this would be a semicore state such as the $3s$ and $3p$ orbitals in titanium. The minimum energy is set below this bound state, $\varepsilon_{min} = \varepsilon_v - \varepsilon_{pad}$, where the padding energy is 0.3 Ha. by default. The energy maximum is set to cover the relevant energy ranges, 50 eV to 100 eV for x-ray absorption transition matrix elements or 100 eV to 200 eV for RPA screening calculations (our default value is 5 Ha. ≈ 130 eV). Strictly speaking, this range depends on the Fermi energy and band gap, but for condensed-matter systems these values only vary by a few eV.

Having defined the system's effective Hamiltonians H^{ae} and H^{ps} and an energy window, we can begin to create the projectors. First, a set of pseudopotential partial waves are created for 128 energies spanning from ε_{min} to ε_{max}

$$H^{ps}\phi_i = \varepsilon_i\phi_i^{ps}. \quad \text{(C1)}$$

Note that the calculation is only carried out to a finite radius, and therefore there is no problem normalizing these states. Additionally, these states are not orthogonal, but instead provide an overdetermined basis.

Next, for each pseudopotential partial wave ϕ_i^{ps} an all-electron partial wave ϕ_i^{ae} is also constructed. The ϕ^{ae} are not constructed to match exactly the energies of their corresponding pseudopotential partial wave, but instead to match the pseudopotential wave function and scattering properties. Specifically, we match the arctangent of the log derivatives of the partial waves β, evaluated at the augmentation radius r_a

$$\beta = \left.\frac{r}{\phi}\frac{d\phi}{dr}\right|_{r=r_a}$$
$$\bar{\delta} = \arctan[\beta] - \pi\eta, \quad \text{(C2)}$$

where η is the number of nodes in the partial wave, corrected for the lack of core-level resonances in the pseudopotential system. We will refer to $\bar{\delta}$ as the phase shift. The true phase shift can be related more carefully to the logarithmic derivative and the partial wave (see chapter 7 in Sakurai [83] among others). The pseudopotential properties, matching energy and smoothly matching the wave functions between the pseudo and all-electron systems, are only exact at specific energies. At other energies the mapping is only approximate, and we chose to enforce smoothness at the expense of the energy. As we are only interested in the spatial behavior of the augmented orbitals, this choice is natural.

Because we are matching phase shifts the energy of the all-electron partial wave is only approximately the same as that of the pseudopotential partial wave. A reference set of all-electron partial waves are constructed within the same energy window. Then the energy of each all-electron partial wave is iteratively refined until the phase shifts converge within 3×10^{-14}. Last, ϕ_i^{ae} is rescaled in a fashion that avoids numerical difficulties in cases of nodes and antinodes approaching r_a for a given energy:

$$\phi_i^{ae} := A_i\phi_i^{ae}$$
$$A_i = \left.\frac{\left(\phi_i^{ps}\right)^2 + \left(d\phi_i^{ps}/dr\right)^2}{\phi_i^{ps}\phi_i^{ae} + \left(d\phi_i^{ps}/dr\right)\left(d\phi_i^{ae}/dr\right)}\right|_{r=r_a}. \quad \text{(C3)}$$

Here the partial waves and derivatives are evaluated at the augmentation radius r_a. In the case where the partial waves and first derivatives are equal we have $A = 1$, whereas we typically find $0.95 \leqslant A \leqslant 1.05$.

We now have a set of all-electron and pseudopotential partial waves. To generate the optimal projectors we use principal-component analysis (PCA) [84]. We generate eigenvectors and eigenvalues of the overlap matrix S, with a matrix element and the kth eigenvalue and eigenvector denoted by

$$S_{ij} = \left(\phi_i^{ps}\right)^\dagger\phi_j^{ps}; \quad Sx_k = e_kx_k. \quad \text{(C4)}$$

Using normalized partial waves the trace of S is the number of projectors $N \leqslant 128$. The eigenvectors are sorted by eigenvalues, and only a few with the largest eigenvalues are kept such that $\sum_k^n e_k > N(1 - \iota)$, where n between 3 and 5 is usually
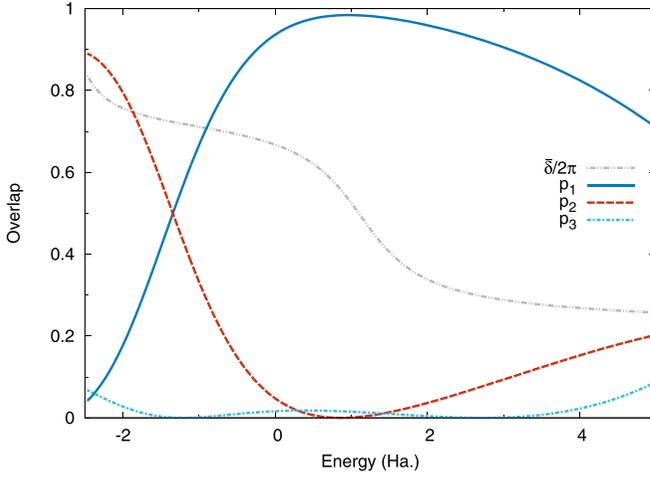
FIG. 15. Overlaps between $l = 0$ pseudopotential scattering states of titanium and the calculated OPFs as a function of the energy of the scattering state. For reference, the rescaled phase shift $\bar{\delta}/2\pi$ is plotted as well. Note the OPFs are not necessarily eigenstates for any energy, i.e., the overlap for any single projector may be always less than 1. This is in contrast to pseudopotential or PAW methods that choose projectors that are eigenstates.

sufficient for a small error ($\iota < 10^{-4}$). We can now construct the optimal projectors following the prescription,

$$p_k^{ae} = \sum_{i=1}^{N} x_{ki} \phi_k^{ae}; \quad p_k^{ps} = \sum_{i=1}^{N} x_{ki} \phi_k^{ps}. \quad (C5)$$

Here both the pseudo and all-electron projectors are constructed from the same vector x_k. (In practice, we build the negative of S so that the eigenvalues with the largest absolute magnitude are also the lowest. We then only calculate 16 eigenvectors with the smallest eigenvalues using the SYEVR routine provided by the LAPACK library [85,86].)

We call these projectors optimal because the PCA construction guarantees that the fewest projectors possible are chosen to span the space given by our set of partial waves and target error. The strength of this approach is that relatively few projectors per angular momentum are needed to span from the occupied valence bands through 130 eV above the Fermi level. The number of projectors generated is typically one more than the number of scattering resonances of spanned by the OPF energy window. In Fig. 15 we show the overlaps between the partial waves and the optimal projectors $|\langle \phi_j^{ps}|p_i^{ps}\rangle|^2$ as a function of energy for the $l = 0$ states of titanium.

The augmentation of the electron wave functions is carried out using the projectors from Eq. (C5). An all-electron wave function is given as follows (here μ denotes, say, a Bloch state, and the atom's position is taken as the origin):

$$\psi_\mu^{ae}(\mathbf{r}) = \psi_\mu^{ps}(\mathbf{r}) + \sum_{lm} \sum_j^{N_l} Y_{lm}(\hat{r})\big[p_{jl}^{ae}(r) - p_{jl}^{ps}(r)\big]$$

$$\times \big\langle Y_{lm} p_{jl}^{ps}|\psi_\mu^{ps}\big\rangle_{r_a}, \quad (C6)$$

where $-l \leqslant m \leqslant +l$, N_l is the number of projectors for a given angular momentum channel, and Y_{lm} are the spherical harmonics. The overlap between the wave function and the projectors is taken within the sphere defined by the with radius r_a.

APPENDIX D: GRIDS, CONVERGENCE, AND ERRORS FROM APPROXIMATIONS

1. Grids and integrals

As outlined previously in Sec. VI on the implementation within OCEAN, the Green's functions are calculated on a real-space grid determined by separate radial and angular grids, and the internal energy loop integral is calculated for a set of imaginary energies.

a. Radial and angular grids

The real-space points used for calculating the Green's functions and polarizability are constructed from separate radial and angular grids,

$$\mathbf{r}_i = r_j \otimes \hat{\Omega}_k. \quad (D1)$$

The angular grids are taken from the set of extremal points by Womersley and Sloan [87]. For a given degree n each angular grid has dimension $(n + 1)^2$. The radial grid has two options, uniform spacing or Gauss-Legendre quadrature. Uniform spacing has the advantage that it is directly equivalent to a plane-wave energy cutoff $|\mathbf{G}_{max}| = \pi/\Delta r$. However, testing indicates that the quadrature grids are more efficient, generating converged results with fewer points. The radial space can divided into arbitrary parts, each with its own grid spacing or quadrature grid.

By default, we divide the space in two using r_a, the augmentation radius from the OPFs. Within this radius we use a 16-point Gauss-Legendre radial grid and the 36-point ($n = 5$) angular grid. The dense radial grid captures the behavior of the all-electron reconstructed wave functions close to the atomic site. For the section outside r_a, we use a uniform grid such that $(r_{max} - r_a)/N < 0.45$ a.u., typically 16 points, and the angular grid is increased to 64 points ($n = 7$). This gives the Green's functions and polarizability dimensions of 1600×1600, independent of the size of the unit cell.

b. Energy integration

By construction, the RPA polarizability requires an integral over the internal loop energy. As shown in Eq. (25) this can be transformed from an integral over real energies to one over complex energies by closing the contour and realizing that above the Fermi energy all of the poles (single-particle excitation energies) are displaced below the real axis by a small imaginary component. Likewise, below the Fermi level the poles are displaced above the real axis. Therefore, the contour is closed by arcs in the upper-right and lower-left quadrants and does not encompass any poles. The Green's function is relatively smooth at imaginary values, and we use quadrature to replace the integral with a summation over relatively few energy points.

Following Ref. [11], we first divide $t \in (-\infty, \infty)$ into four regions, symmetric across $t = 0$ by the parameter ζ. such that the number of quadrature points in the region $(0, \zeta)$ will be the

same as within (ζ, ∞). This allows Eq. (25) to be rewritten as

$$\chi_0 = \sum_\sigma \int_{-\infty}^\infty \frac{dt}{2\pi} g_\sigma(\mathbf{r}, \mathbf{r}', \mu + it) g_\sigma(\mathbf{r}', \mathbf{r}, \mu + it)$$

$$= \sum_\sigma \frac{\zeta}{\pi} \int_0^1 da [g_\sigma(\mathbf{r}, \mathbf{r}', \mu + i\zeta a) g_\sigma(\mathbf{r}', \mathbf{r}, \mu + i\zeta a)$$

$$+ a^{-2} g_\sigma(\mathbf{r}, \mathbf{r}', \mu + {}^{i\zeta}/a) g_\sigma(\mathbf{r}', \mathbf{r}, \mu + {}^{i\zeta}/a)]. \quad (D2)$$

The parameter ζ is chosen to be the geometric mean of the largest and smallest values of $|\mu - \varepsilon_{b\mathbf{k}}|$, i.e., half the band gap and the larger of the distance from μ to the bottom of the valence bands or the top of the conductions bands. To prevent ζ from going to zero in metallic systems, 0.5 eV is added in quadrature to the minimum (half-gap) value.

The integral over a in Eq. (D2) is replaced by a summation over quadrature points. The energy points a_i and weights w_i are taken from Gauss-Legendre quadrature, shifted and scaled by half to match the range. Quadrature grids from 4 to 64 points are implemented in the code. In Ref. [11], it was suggested that the two-part integrand be replaced with a single product of Green's functions with energy points at $i\zeta a/(1 - a)$ with a prefactor of $(1 - a)^{-2}$, giving

$$\chi_0 \approx \sum_\sigma \sum_i^{N_i} \frac{\zeta w_i}{\pi (1 - a_i)^2} g_\sigma\left(\mathbf{r}, \mathbf{r}', \mu + i\frac{\zeta a_i}{1 - a_i}\right)$$

$$\times g_\sigma\left(\mathbf{r}', \mathbf{r}, \mu + i\frac{\zeta a_i}{1 - a_i}\right). \quad (D3)$$

This reproduces the correct large and small a behavior of Eq. (D2), but with only a single set of quadrature points. Using this single-grid approximation, a 16-point quadrature grid was found to be sufficient. For systems with time-reversal symmetry, the spatial indices on one of the Green's functions can be interchanged. This allows us to calculate only a single Green's function and square it.

2. Bands and k-points

The convergence of the screening calculation also depends on the number of k-points and bands included in the Green's functions. The convergence behavior with respect to bands is similar in our approach and other sum over states methods. A large number of unoccupied bands may be required, and the error falls as the inverse of the number of bands [88]. To generalize between materials it is preferable to speak of the energy range of the unoccupied bands included in the calculation, e.g., the average energy of the highest-band with respect to the conduction band minimum. In Fig. 16(b) we show the convergence of the screening potential of the fluorine $1s$ hole in LiF with respect to the energy range of unoccupied states. This is done by plotting the difference between the calculated induced potential using a conduction band range of 200 eV and that calculated with smaller ranges, e.g., $\Delta v_{\text{ind}}[100 \text{ eV}] = v_{\text{ind}}[200 \text{ eV}] - v_{\text{ind}}[100 \text{ eV}]$, etc. Typically, the induced potential near the core hole increases with an increase in the number of bands included in the screening calculation.

Like the summation over bands, the k-point sampling should also be infinite. However, while the summation over bands takes the place of an energy integral whose upper
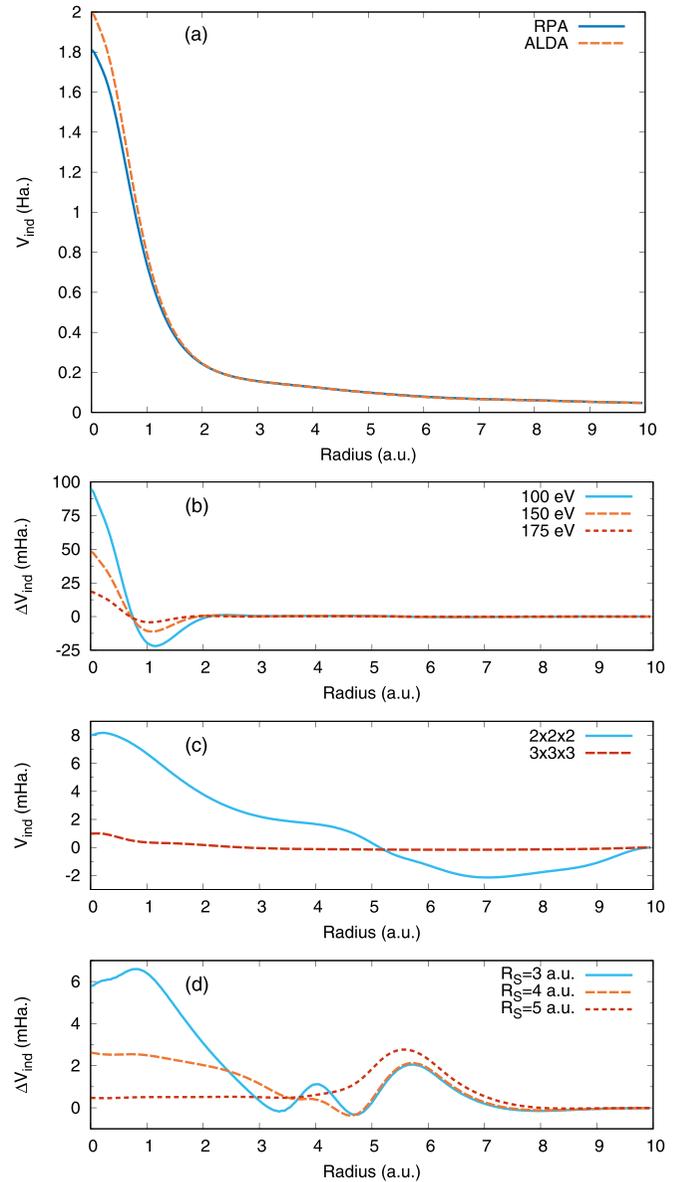

FIG. 16. (a) The induced potential in response to a fluorine $1s$ hole in LiF calculated using a shell radius of 6 a.u. and 120 conduction bands, spanning approximately 200 eV, and a 4^3 k-point sampling (solid, blue). The orange, dashed line shows the induced potential calculated using f_{xc} within the adiabatic LDA (see text Sec. VII C). (b) The difference plots obtained by subtracting the induced potential calculated with different numbers of bands. (c) The difference plots for changing k-point grids. (d) The difference plots for the induced potential changing only the sphere radius R_S. Note that the difference plots are in mHa.

bound is positive infinity, the summation over k-points is, by construction, a properly normalized discretization of the volume integral over the Brillouin zone. Errors in finite k-point sampling arise when the electron wave functions at a given momentum are a poor approximation of other points within the discretization volume. In the real-space approach, the convergence with k-points is rapid. Even for systems with small units cells like LiF, only a few k-points are required. In Fig. 16(c) we show the difference plots from reducing the
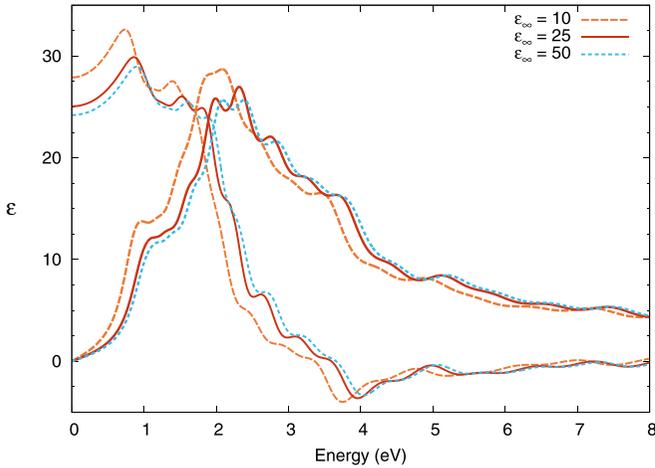
FIG. 17. The real (thin) and imaginary (thick) components of the complex dielectric constant of FeS_2 plotted for 3 different inputs of the static dielectric constant ϵ_∞. If the value of ϵ_∞ is not known from prior calculations or experiment, then valence BSE calculations can be carried out to determine it self-consistently: $\epsilon_\infty = 25$.

k-point sampling from 4^3 down to 3^3 and 2^3. With only a $2 \times 2 \times 2$ k-point grid, the errors in the induced potential are less than 10 mHa.

3. Real-space truncation

As introduced in Sec. II C, the real-space approach relies on partitioning the space around the core hole (or test charge). This partitioning is carried out in Eq. (9), where a spherical charge of radius R_S neutralizes the long-range Coulomb tail, allowing the RPA screening to be carried out only locally. Our approximation doesn't change the total external potential that is screened. However, by using a model dielectric to calculate $W^{(2)}$ we introduce differences with respect to a calculation using the RPA polarizability everywhere. Having previously outlined the effect of neglecting the core-hole potential and the effects of various approximations to the augmenting the pseudopotential wave functions in the previous section, we now look to the influence of R_S on the calculated screened core-hole potential and subsequently the absorption spectrum.

To assess the effects of finite R_S on the calculated screened potential we compare the induced potential for the fluorine K edge of LiF. In Fig. 16(d) we show the induced potential calculated with a shell radius of 6 a.u. and the difference in the potentials between those calculated with shell radii of 5, 4, and 3 a.u: $\Delta v_{ind}[R_S = 5] = v_{ind}[R_S = 6] - v_{ind}[R_S = 5]$. Near the fluorine site, the difference between the induced potential at $R_S = 6$ a.u. and $R_S = 5$ a.u. is less than 0.013 eV, while the maximum difference between the two, located at 5.55 a.u. (approximately the length of a lattice vector), is less than 0.076 eV.

4. Long-range dielectric constant

As mentioned previously, the static long-range dielectric constant ϵ_∞ is a required input for our local, real-space approach. The error near the core hole due to an incorrect
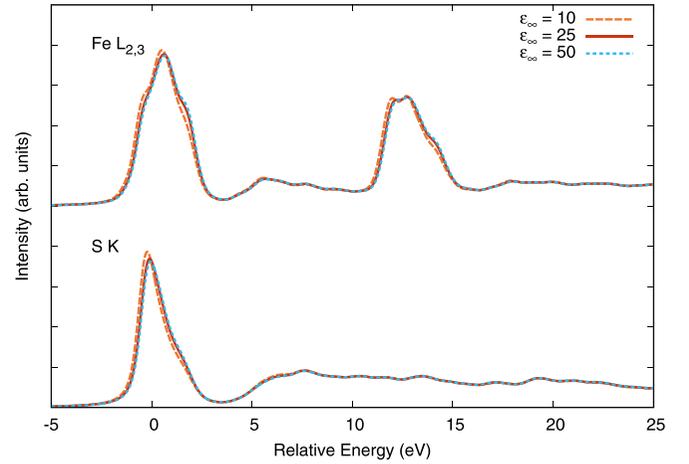

FIG. 18. For FeS_2 marcasite, the Fe $L_{2,3}$ edge and S K edges are plotted for three different inputs of the static dielectric constant ϵ_∞. Despite large differences in the input ϵ_∞, the spectra are largely unchanged. The relative positioning takes into account the changes in the effect of the changing screened core-hole potential on the chemical shifts (see text). The error in the position of either edge is less than 0.13 eV for $\epsilon_\infty = 10$.

dielectric constant is approximately

$$\Delta W \approx \left(1 - \frac{1}{\epsilon_\infty}\right)R_S^{-1} - \left(1 - \frac{1}{\widetilde{\epsilon_\infty}}\right)R_S^{-1}, \quad \text{(D4)}$$

where $\widetilde{\epsilon_\infty}$ is the input dielectric constant. This can be expressed in terms of percentage error in the input dielectric constant

$$\Delta W \approx \frac{1}{\widetilde{\epsilon_\infty}R_S}\left(\frac{\epsilon_\infty - \widetilde{\epsilon_\infty}}{\epsilon_\infty}\right). \quad \text{(D5)}$$

As an example, for $\epsilon_\infty = 5$, $R_S = 5$ a.u., a 10% underestimation ($\widetilde{\epsilon_\infty} = 4.5$) would lead to an error of 0.12 eV. This absolute error directly affects calculations of chemical shifts using the OCEAN code [29].

To showcase the errors from an incorrect input value of ϵ_∞ we look at FeS_2 in the cubic *Pnnm* phase marcasite. As before the cell parameters are taken from experiment [89], the pseudopotentials were taken from PseduoDojo, specifically the "high-accuracy" iron potential. Marcasite crystalizes in the cubic *Pnnm* phase. The lattice parameters were set to 0.44446 by 0.54246 nm by 0.33864 nm to match experiment [89]. A plane-wave cutoff of 100 Ry was used and the density was converged on a 4^3 k-point mesh. The "high-accuracy" iron and standard sulfur pseudopotentials were taken from PseudoDojo. The BSE final states were solved on a $12 \times 10 \times 16$ k-point mesh, including 72 unoccupied bands, and downsampled onto an $8 \times 10 \times 6$ real-space mesh. For the screening calculations a 2^3 k-point mesh and 200 bands were used. Absent a previously calculated or experimentally measured value for the static dielectric constant, the input ϵ_∞ can be determined self-consistently as shown in Fig. 17. We find that an input value of $\epsilon_\infty = 25$ results in a BSE calculation of approximately the same value (25.06) with the photon momentum vector aligned along (111). Unsurprisingly, the strength of the static dielectric

constant affects the calculated dielectric response, but even comparing $\epsilon_\infty = 10$ to $\epsilon_\infty = 50$ the spectra are in qualitative agreement.

Next we can compare this effect on the x-ray edges of FeS_2 in Fig. 18. Neither the sulfur K edge nor the iron L edge are strongly dependent on the input value of the dielectric constant. The energy scale of both is relative to the conduction band minimum of the $\epsilon_\infty = 10$ calculation (of the L_3 edge of iron). The slight shifts in the onset of the $\epsilon_\infty = 25$ and $\epsilon_\infty = 50$ spectra are due to changes in the excitonic binding and core-level shift due to differences in the input dielectric. As can be seen in Fig. 18, for x-ray absorption calculations it may be sufficient to have only a rough estimate of the dielectric constant.

[1] N. Wiser, Dielectric constant with local field effects included, Phys. Rev. **129**, 62 (1963).

[2] H. N. Rojas, R. W. Godby, and R. J. Needs, Space-Time Method for *Ab Initio* Calculations of Self-Energies and Dielectric Response Functions of Solids, Phys. Rev. Lett. **74**, 1827 (1995).

[3] X. Blase, A. Rubio, S. G. Louie, and M. L. Cohen, Mixed-space formalism for the dielectric response in periodic systems, Phys. Rev. B **52**, R2225 (1995).

[4] L. Hung, F. H. da Jornada, J. Souto-Casares, J. R. Chelikowsky, S. G. Louie, and S. Öğüt, Excitation spectra of aromatic molecules within a real-space *GW*-BSE formalism: Role of self-consistency and vertex corrections, Phys. Rev. B **94**, 085125 (2016).

[5] F. Giustino, M. L. Cohen, and S. G. Louie, *GW* method with the self-consistent sternheimer equation, Phys. Rev. B **81**, 115105 (2010).

[6] H.-V. Nguyen, T. A. Pham, D. Rocca, and G. Galli, Improving accuracy and efficiency of calculations of photoemission spectra within the many-body perturbation theory, Phys. Rev. B **85**, 081101(R) (2012).

[7] M. D. Ben, F. H. da Jornada, A. Canning, N. Wichmann, K. Raman, R. Sasanka, C. Yang, S. G. Louie, and J. Deslippe, Large-scale *gw* calculations on pre-exascale hpc systems, Comput. Phys. Commun. **235**, 187 (2019).

[8] M. M. Rieger, L. Steinbeck, I. White, H. Rojas, and R. Godby, The gw space-time method for the self-energy of large systems, Comput. Phys. Commun. **117**, 211 (1999).

[9] M. Kaltak, J. Klimeš, and G. Kresse, Cubic scaling algorithm for the random phase approximation: Self-interstitials and vacancies in si, Phys. Rev. B **90**, 054115 (2014).

[10] M. Kim, G. J. Martyna, and S. Ismail-Beigi, Complex-time shredded propagator method for large-scale *gw* calculations, Phys. Rev. B **101**, 035139 (2020).

[11] E. L. Shirley, Local screening of a core hole: A real-space approach applied to hafnium oxide, Ultramicroscopy **106**, 986 (2006).

[12] G. Onida, L. Reining, and A. Rubio, Electronic excitations: Density-functional versus many-body green's-function approaches, Rev. Mod. Phys. **74**, 601 (2002).

[13] Z. H. Levine and S. G. Louie, New model dielectric function and exchange-correlation potential for semiconductors and insulators, Phys. Rev. B **25**, 6310 (1982).

[14] E. L. Shirley, J. A. Soininen, and J. J. Rehr, Modeling core-hole screening in core-excitation spectroscopies, Phys. Scr. **T115**, 31 (2005).

[15] J. Vinson, J. J. Rehr, J. J. Kas, and E. L. Shirley, Bethe-salpeter equation calculations of core excitation spectra, Phys. Rev. B **83**, 115106 (2011).

[16] K. Gilmore, J. Vinson, E. Shirley, D. Prendergast, C. Pemmaraju, J. Kas, F. Vila, and J. Rehr, Efficient implementation of core-excitation bethe-salpeter equation calculations, Comput. Phys. Commun. **197**, 109 (2015).

[17] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, UK, 2004).

[18] L. Kleinman and D. M. Bylander, Efficacious form for Model Pseudopotentials, Phys. Rev. Lett. **48**, 1425 (1982).

[19] E. L. Shirley and R. M. Martin, Gw quasiparticle calculations in atoms, Phys. Rev. B **47**, 15404 (1993).

[20] E. L. Shirley, Bethe-salpeter treatment of x-ray absorption including core-hole multiplet effects, J. Electr. Spectrosc. Relat. Phenom. **144-147**, 1187 (2005).

[21] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

[22] 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. DiStasio, 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, Advanced capabilities for materials modelling with quantum ESPRESSO, J. Phys.: Condens. Matter **29**, 465901 (2017).

[23] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, Quantum espresso: A modular and open-source software project for quantum simulations of materials, J. Phys.: Condens. Matter **21**, 395502 (2009). www.quantum-espresso.org.

[24] J. P. Perdew and Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, Phys. Rev. B **45**, 13244 (1992).

[25] M. van Setten, M. Giantomassi, E. Bousquet, M. Verstraete, D. Hamann, X. Gonze, and G.-M. Rignanese, The pseudodojo: Training and grading a 85 element optimized norm-conserving pseudopotential table, Comput. Phys. Commun. **226**, 39 (2018).

[26] D. R. Hamann, Optimized norm-conserving vanderbilt pseudopotentials, Phys. Rev. B **88**, 085117 (2013).

[27] C. P. Schwartz, F. Ponce, S. Friedrich, S. P. Cramer, J. Vinson, and D. Prendergast, Temperature and radiation effects at the fluorine k-edge in lif, J. Electron Spectrosc. Relat. Phenom. **218**, 30 (2017).

[28] T. A. Pascal, U. Boesenberg, R. Kostecki, T. J. Richardson, T.-C. Weng, D. Sokaras, D. Nordlund, E. McDermott, A. Moewes, J. Cabana, and D. Prendergast, Finite temperature effects on the x-ray absorption spectra of lithium compounds: First-principles interpretation of x-ray raman measurements, J. Chem. Phys. **140**, 034107 (2014).

[29] J. Vinson, T. Jach, W. T. Elam, and J. D. Denlinger, Origins of extreme broadening mechanisms in near-edge x-ray spectra of nitrogen compounds, Phys. Rev. B **90**, 205207 (2014).

[30] C. Brouder, D. Cabaret, A. Juhin, and P. Sainctavit, Effect of atomic vibrations on the x-ray absorption spectra at the k edge of al in α-Al_2O_3 and of ti in TiO_2 rutile, Phys. Rev. B **81**, 115125 (2010).

[31] J. Vinson, T. Jach, M. Müller, R. Unterumsberger, and B. Beckhoff, Resonant x-ray emission of hexagonal boron nitride, Phys. Rev. B **96**, 205116 (2017).

[32] W. Olovsson, T. Mizoguchi, M. Magnuson, S. Kontur, O. Hellman, I. Tanaka, and C. Draxl, Vibrational effects in x-ray absorption spectra of two-dimensional layered materials, J. Phys. Chem. C **123**, 9688 (2019).

[33] N. L. McDougall, J. G. Partridge, R. J. Nicholls, S. P. Russo, and D. G. McCulloch, Influence of point defects on the near edge structure of hexagonal boron nitride, Phys. Rev. B **96**, 144106 (2017).

[34] R. Geick, C. H. Perry, and G. Rupprecht, Normal modes in hexagonal boron nitride, Phys. Rev. **146**, 543 (1966).

[35] A. Gulans, S. Kontur, C. Meisenbichler, D. Nabok, P. Pavone, S. Rigamonti, S. Sagmeister, U. Werner, and C. Draxl, exciting: A full-potential all-electron package implementing density-functional theory and many-body perturbation theory, J. Phys.: Condens. Matter **26**, 363202 (2014).

[36] C. Vorwerk, C. Cocchi, and C. Draxl, Addressing electron-hole correlation in core excitations of solids: An all-electron many-body approach from first principles, Phys. Rev. B **95**, 155121 (2017).

[37] A. J. Morris, R. J. Nicholls, C. J. Pickard, and J. R. Yates, Optados: A tool for obtaining density of states, core-level and optical spectra from electronic structure codes, Comput. Phys. Commun. **185**, 1477 (2014).

[38] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, First principles methods using CASTEP, Z. Kristallogr. **220**, 567 (2005).

[39] A. Preobrajenski, A. Vinogradov, and N. Mårtensson, Decay of core excitations in bulk h-BN studied with resonant auger spectroscopy, J. Electr. Spectrosc. Relat. Phenom. **148**, 59 (2005).

[40] N. L. McDougall, R. J. Nicholls, J. G. Partridge, and D. G. McCulloch, The near edge structure of hexagonal boron nitride, Microsc. Microanal. **20**, 1053 (2014).

[41] Y. Liang, J. Vinson, S. Pemmaraju, W. S. Drisdell, E. L. Shirley, and D. Prendergast, Accurate X-Ray Spectral Predictions: An Advanced Self-Consistent-Field Approach Inspired by Many-Body Perturbation Theory, Phys. Rev. Lett. **118**, 096402 (2017).

[42] L. X. Benedict and E. L. Shirley, Ab initio calculation of $\epsilon_2(\omega)$ including the electron-hole interaction: Application to gan and CaF_2, Phys. Rev. B **59**, 5441 (1999).

[43] H. M. Lawler, J. J. Rehr, F. Vila, S. D. Dalosto, E. L. Shirley, and Z. H. Levine, Optical to UV spectra and birefringence of SiO_2 and TiO_2: First-principles calculations with excitonic effects, Phys. Rev. B **78**, 205108 (2008).

[44] M. S. Hybertsen and S. G. Louie, Model dielectric matrices for quasiparticle self-energy calculations, Phys. Rev. B **37**, 2733 (1988).

[45] P. Lautenschlager, M. Garriga, L. Vina, and M. Cardona, Temperature dependence of the dielectric function and interband critical points in silicon, Phys. Rev. B **36**, 4821 (1987).

[46] D. M. Roessler and W. C. Walker, Optical constants of magnesium oxide and lithium fluoride in the far ultraviolet, J. Opt. Soc. Am. **57**, 835 (1967).

[47] W. Hanke and L. J. Sham, Many-Particle Effects in the Optical Excitations of a Semiconductor, Phys. Rev. Lett. **43**, 387 (1979).

[48] R. W. G. Wyckoff, *Crystal Structures*, 2nd ed. (Interscience, New York, 1963).

[49] C. Kittel, *Introduction to Solid State Physics*, 8th ed. (John Wiley & Sons, New York, 2005).

[50] P. Puschnig and C. Ambrosch-Draxl, Optical absorption spectra of semiconductors and insulators including electron-hole correlations: An *ab initio* study within the lapw method, Phys. Rev. B **66**, 165105 (2002).

[51] M. Gatti and F. Sottile, Exciton dispersion from first principles, Phys. Rev. B **88**, 155113 (2013).

[52] M. Piacentini, D. W. Lynch, and C. G. Olson, Thermoreflectance of LiF between 12 and 30 eV, Phys. Rev. B **13**, 5530 (1976).

[53] 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, The abinit project: Impact, environment and recent developments, Comput. Phys. Commun. **248**, 107042 (2020).

[54] 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, Many-body perturbation theory calculations using the yambo code, J. Phys.: Condens. Matter **31**, 325902 (2019).

[55] X. Zhu and S. G. Louie, Quasiparticle band structure of thirteen semiconductors and insulators, Phys. Rev. B **43**, 14142 (1991).

[56] A form of this has been attributed to Enrico Fermi.

[57] A. Ware, Fast approximate fourier transforms for irregularly spaced data, SIAM Rev. **40**, 838 (1998).

[58] Specific hardware is named for identification purposes; it does not imply recommendation or endorsement by NIST. The CPUs used were Intel Xeon E5-2630 v3.

[59] F. Bruneval and X. Gonze, Accurate *GW* self-energies in a plane-wave basis using only a few empty states: Towards large systems, Phys. Rev. B **78**, 085125 (2008).

[60] J. A. Berger, L. Reining, and F. Sottile, *Ab initio* calculations of electronic excitations: Collapsing spectral sums, Phys. Rev. B **82**, 041103(R) (2010).

[61] J. A. Berger, L. Reining, and F. Sottile, Efficient *GW* calculations for SnO_2, ZnO, and rubrene: The effective-energy technique, Phys. Rev. B **85**, 085126 (2012).

[62] R. James and S. Woodley, Extraction of green's functions from total energy plane wave pseudopotential calculations, Solid State Commun. **97**, 935 (1996).

[63] L. Steinbeck, A. Rubio, L. Reining, M. Torrent, I. White, and R. Godby, Enhancements to the GW space-time method, Comput. Phys. Commun. **125**, 105 (2000).

[64] G. Samsonidze, M. Jain, J. Deslippe, M. L. Cohen, and S. G. Louie, Simple Approximate Physical Orbitals for *GW* Quasiparticle Calculations, Phys. Rev. Lett. **107**, 186404 (2011).

[65] J. Klimeš, M. Kaltak, and G. Kresse, Predictive *GW* calculations using plane waves and pseudopotentials, Phys. Rev. B **90**, 075125 (2014).

[66] H. Lambert and F. Giustino, *Ab initio* sternheimer-gw method for quasiparticle calculations using plane waves, Phys. Rev. B **88**, 075117 (2013).

[67] H. F. Wilson, F. Gygi, and G. Galli, Efficient iterative method for calculations of dielectric matrices, Phys. Rev. B **78**, 113303 (2008).

[68] L. Hedin, New method for calculating the one-particle green's function with application to the electron-gas problem, Phys. Rev. **139**, A796 (1965).

[69] W. Gao and J. R. Chelikowsky, Real-space based benchmark of g0w0 calculations on GW100: Effects of semicore orbitals and orbital reordering, J. Chem. Theory Comput. **15**, 5299 (2019).

[70] B.-C. Shih, Y. Xue, P. Zhang, M. L. Cohen, and S. G. Louie, Quasiparticle Band Gap of ZnO: High Accuracy from the Conventional G^0W^0 Approach, Phys. Rev. Lett. **105**, 146401 (2010).

[71] D. Golze, J. Wilhelm, M. J. van Setten, and P. Rinke, Core-level binding energies from GW: An efficient full-frequency approach within a localized basis, J. Chem. Theory Comput. **14**, 4856 (2018).

[72] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Phonons and related crystal properties from density-functional perturbation theory, Rev. Mod. Phys. **73**, 515 (2001).

[73] X. Gonze, First-principles responses of solids to atomic displacements and homogeneous electric fields: Implementation of a conjugate-gradient algorithm, Phys. Rev. B **55**, 10337 (1997).

[74] X. Gonze and C. Lee, Dynamical matrices, born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory, Phys. Rev. B **55**, 10355 (1997).

[75] A. A. Quong and B. M. Klein, Self-consistent-screening calculation of interatomic force constants and phonon dispersion curves from first principles, Phys. Rev. B **46**, 10734 (1992).

[76] R. Del Sole, L. Reining, and R. W. Godby, $GW\Gamma$ approximation for electron self-energies in semiconductors and insulators, Phys. Rev. B **49**, 8024 (1994).

[77] A. J. Morris, M. Stankovski, K. T. Delaney, P. Rinke, P. García-González, and R. W. Godby, Vertex corrections in localized and extended systems, Phys. Rev. B **76**, 155106 (2007).

[78] F. Bruneval, F. Sottile, V. Olevano, R. Del Sole, and L. Reining, Many-Body Perturbation Theory using the Density-Functional Concept: Beyond the *GW* Approximation, Phys. Rev. Lett. **94**, 186402 (2005).

[79] M. L. Tiago and J. R. Chelikowsky, Optical excitations in organic molecules, clusters, and defects studied by first-principles green's function methods, Phys. Rev. B **73**, 205334 (2006).

[80] J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, Phys. Rev. B **23**, 5048 (1981).

[81] S. H. Vosko, L. Wilk, and M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis, Can. J. Phys. **58**, 1200 (1980).

[82] D. M. Ceperley and B. J. Alder, Ground State of the Electron Gas by a Stochastic Method, Phys. Rev. Lett. **45**, 566 (1980).

[83] J. J. Sakurai, *Modern Quantum Mechanics*, edited by S. F. Tuan (Addison-Wesley, Reading, MA, 1994).

[84] K. Pearson, On lines and planes of closest fit to systems of points in space, Lond. Edinb. Dubl. Philos. Mag. J. Sci. **2**, 559 (1901).

[85] I. S. Dhillon and B. N. Parlett, Multiple representations to compute orthogonal eigenvectors of symmetric tridiagonal matrices, Linear Algebr. Appl. **387**, 1 (2004).

[86] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*, 3rd ed. (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999).

[87] I. H. Sloan and R. S. Womersley, Extremal systems of points and numerical integration on the sphere, Adv. Comput. Math. **21**, 107 (2004).

[88] C. Friedrich, M. C. Müller, and S. Blügel, Band convergence and linearization error correction of all-electron *GW* calculations: The extreme case of zinc oxide, Phys. Rev. B **83**, 081101(R) (2011).

[89] M. Rieder, J. C. Crelling, O. Šustai, M. Drábek, Z. Weiss, and M. Klementová, Arsenic in Iron Disulfides in a Brown Coal from the North Bohemian Basin, Czech Republic, Int. J. Coal Geol. **71**, 115 (2007).