

All-electron Gaussian-based G_0W_0 for Valence and Core Excitation Energies of Periodic Systems

Tianyu Zhu* and Garnet Kin-Lic Chan†

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena CA 91125

We describe an all-electron G_0W_0 implementation for periodic systems with \mathbf{k} -point sampling implemented in a crystalline Gaussian basis. Our full-frequency G_0W_0 method relies on efficient Gaussian density fitting integrals and includes both analytic continuation and contour deformation schemes. Due to the compactness of Gaussian bases, no virtual state truncation is required as is seen in many plane-wave formulations. Finite size corrections are included by taking the $\mathbf{q} \rightarrow \mathbf{0}$ limit of the Coulomb divergence. Using our implementation, we study quasiparticle energies and band structures across a range of systems including molecules, semiconductors, rare gas solids, and metals. We find that the G_0W_0 band gaps of traditional semiconductors converge rapidly with respect to the basis size, even for the conventionally challenging case of ZnO. Using correlation-consistent bases of polarized triple-zeta quality, we find the mean absolute relative error of the extrapolated G_0W_0 @PBE band gaps to be only 5.2% when compared to experimental values. For core excitation binding energies (CEBEs), we find that G_0W_0 predictions improve significantly over those from DFT if the G_0W_0 calculations are started from hybrid functionals with a high percentage of exact exchange.

I. INTRODUCTION

The accurate simulation of materials spectra is an important target of computational solid-state physics and chemistry. While density functional theory (DFT) [1] has been widely used due to its low computational cost, the derivative discontinuity means that Kohn-Sham orbital energies do not formally describe the quasiparticle energies [2–4]. In practice, band structures obtained from the local density approximation (LDA) and generalized gradient approximations (GGA) substantially underestimate band gaps in most solids [5]. A many-body computation of the quasiparticle energies is thus desirable, and the GW approximation [6–9] is often used for this task. The GW approximation starts from a screened Coulomb interaction (W) treated at the level of the random phase approximation (RPA) [10], and the method derives its name from defining the self-energy as $\Sigma = G * W$, where G is the one-electron Green’s function. Because of the simple structure of the self-energy and the associated low computational cost compared to other many-body treatments, GW has become a method of choice for quantitative simulations of quasiparticle spectra in weakly-correlated materials.

Within the GW approximation, one can define several levels of self-consistency. Often, GW is employed in its cheapest form, as a non-self-consistent, one-shot, method referred to as G_0W_0 . In this case, the accuracy of the method has a dependence on the starting choice of orbitals and orbital energies which enter into G_0 . Like other electronic structure methods, how best to implement G_0W_0 depends on the computational basis. The most common basis with which to implement G_0W_0 in periodic systems is the plane-wave (PW) basis [11–15]. However, this basis has some well-known drawbacks. One is that representing the sharp core electron density well requires a large number of plane waves. Thus either pseudopotentials must be used, or the basis should be augmented by other functions, as is done in muffin tin representations, such as the full-potential linear augmented plane waves

(FLAPW) [16–18], linear muffin-tin orbitals (LMTO) [19], and projector augmented-wave (PAW) [20–23] techniques. Pseudopotentials prevent access to the core excitation binding energies (CEBEs), and while augmented plane wave (APW) schemes in principle allow core states to be accessed, the implementations are more complicated [24, 25]. A second drawback of plane-wave based G_0W_0 implementations is the need for a large number of virtual states to converge the screened interaction (even when computing valence quasiparticle energies [9]), especially for materials with d or f electrons [26–28]. Some techniques to alleviate this problem by avoiding explicit sums over virtual states [13, 15, 29] have recently appeared.

Local atomic orbital representations (AOs), such as the Gaussian basis sets that are widely used in quantum chemistry methods, are another potential choice of computational basis. Local AOs are well suited to describe the rapid oscillations of the electron density near nuclei, and treat the core and valence states on an equal footing. Furthermore, due to the much smaller basis size of typical AO basis sets compared to plane wave bases [30], summations over virtual states can be directly performed without truncation. Several G_0W_0 implementations based on local AO bases, including Gaussians [31–35] and numeric atom-centered orbitals (NAOs) [36–38], have been reported for molecules. However, there are few such implementations for periodic systems. Amongst these, Rohlfing *et al.* carried out some early explorations using Gaussian orbitals for G_0W_0 within the plasmon-pole model [39, 40], and recently a Γ -point only periodic G_0W_0 implementation using Gaussian basis sets in conjunction with pseudopotentials has been reported [41].

In this work, we describe a periodic all-electron full-frequency G_0W_0 implementation using (crystalline) Gaussian basis sets within the PySCF quantum chemistry platform [42, 43]. We note that this implementation was previously used, without a description of the implementation or its performance, in the context of the full cell GW +DMFT approach that we have recently developed for strongly correlated solids [44–47]. The purpose of this work is to carefully present the algorithm and benchmark its performance. To this end, we carry out detailed benchmarks for quasiparticle energies, pay-

* tyzhu@caltech.edu

† gkc1000@gmail.com

ing attention to basis set convergence and finite size effects, across a wide range of problems including the excitation energies of molecules, semiconductors, rare gas solids, and metals, and core excitation binding energies in semiconductors. Our G_0W_0 implementation explicitly samples the Brillouin zone (i.e., uses \mathbf{k} -points), and utilizes the efficient periodic Gaussian density fitting integral infrastructure in PySCF [48]. Different numerical treatments are appropriate for the core and valence excitations, and we describe and contrast implementations based on the analytic continuation (AC) [16, 21, 49, 50] and contour deformation (CD) [13, 22, 51] techniques.

II. THEORY AND IMPLEMENTATION

A. G_0W_0 Approximation

Our periodic G_0W_0 implementation closely follows the molecular G_0W_0 implementations described in Refs. [34, 36, 37]. Assuming a DFT calculation has already been performed on a given periodic system, the key quantity to compute in G_0W_0 is the self-energy

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{2\pi} \int_{-\infty}^{\infty} d\omega' e^{i\omega'\eta} G_0(\mathbf{r}, \mathbf{r}', \omega + \omega') W_0(\mathbf{r}, \mathbf{r}', \omega'). \quad (1)$$

Here, G_0 is the non-interacting Green's function, W_0 is the screened Coulomb interaction, ω is the frequency and η is a positive infinitesimal. $G_0(\mathbf{r}, \mathbf{r}', \omega)$ is defined from the DFT (crystalline) molecular orbital (MO) energies $\{\epsilon_{m\mathbf{k}_m}\}$ and orbitals $\{\psi_{m\mathbf{k}_m}\}$:

$$G_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_{m\mathbf{k}_m} \frac{\psi_{m\mathbf{k}_m}(\mathbf{r})\psi_{m\mathbf{k}_m}^*(\mathbf{r}')}{\omega - \epsilon_{m\mathbf{k}_m} - i\eta \text{sgn}(\epsilon_F - \epsilon_{m\mathbf{k}_m})}, \quad (2)$$

with ϵ_F as the Fermi energy and \mathbf{k}_m as a crystal momentum vector in the first Brillouin zone. The screened Coulomb interaction $W_0(\mathbf{r}, \mathbf{r}', \omega)$ is defined as

$$W_0(\mathbf{r}, \mathbf{r}', \omega) = \int d\mathbf{r}'' \epsilon^{-1}(\mathbf{r}, \mathbf{r}'', \omega) v(\mathbf{r}'', \mathbf{r}'), \quad (3)$$

where $v(\mathbf{r}, \mathbf{r}') = |\mathbf{r} - \mathbf{r}'|^{-1}$ is the bare Coulomb operator, and $\epsilon(\mathbf{r}, \mathbf{r}', \omega)$ is the dielectric function:

$$\epsilon(\mathbf{r}, \mathbf{r}', \omega) = \delta(\mathbf{r}, \mathbf{r}') - \int d\mathbf{r}'' v(\mathbf{r}, \mathbf{r}'') \chi_0(\mathbf{r}'', \mathbf{r}', \omega). \quad (4)$$

The polarizability $\chi_0(\mathbf{r}, \mathbf{r}', \omega)$ is calculated in the random phase approximation (RPA):

$$\chi_0(\mathbf{r}, \mathbf{r}', \omega) = \frac{1}{N_{\mathbf{k}}} \sum_{i\mathbf{k}_i}^{\text{occ}} \sum_{a\mathbf{k}_a}^{\text{vir}} \left(\frac{\psi_{i\mathbf{k}_i}^*(\mathbf{r})\psi_{a\mathbf{k}_a}(\mathbf{r})\psi_{a\mathbf{k}_a}^*(\mathbf{r}')\psi_{i\mathbf{k}_i}(\mathbf{r}')}{\omega - (\epsilon_{a\mathbf{k}_a} - \epsilon_{i\mathbf{k}_i}) + i\eta} - \frac{\psi_{i\mathbf{k}_i}(\mathbf{r})\psi_{a\mathbf{k}_a}^*(\mathbf{r})\psi_{a\mathbf{k}_a}(\mathbf{r}')\psi_{i\mathbf{k}_i}^*(\mathbf{r}')}{\omega + (\epsilon_{a\mathbf{k}_a} - \epsilon_{i\mathbf{k}_i}) - i\eta} \right), \quad (5)$$

where i and a label occupied and virtual molecular orbitals respectively.

Once the self-energy is computed, one solves the G_0W_0 quasiparticle (QP) equation to obtain G_0W_0 QP energies:

$$\epsilon_{n\mathbf{k}}^{GW} = \epsilon_{n\mathbf{k}}^{\text{DFT}} + (\psi_{n\mathbf{k}} | \text{Re} \Sigma(\epsilon_{n\mathbf{k}}^{GW}) - v^{xc} | \psi_{n\mathbf{k}}). \quad (6)$$

Eq. 6 needs to be solved self-consistently. Sometimes an approximate linearization is used:

$$\epsilon_{n\mathbf{k}}^{GW} = \epsilon_{n\mathbf{k}}^{\text{DFT}} + Z_{n\mathbf{k}} (\psi_{n\mathbf{k}} | \text{Re} \Sigma(\epsilon_{n\mathbf{k}}^{\text{DFT}}) - v^{xc} | \psi_{n\mathbf{k}}), \quad (7)$$

where $Z_{n\mathbf{k}}$ is the renormalization factor

$$Z_{n\mathbf{k}} = \left(1 - \left. \frac{\partial \text{Re} \Sigma_{n\mathbf{k}}(\omega)}{\partial \omega} \right|_{\omega=\epsilon_{n\mathbf{k}}} \right)^{-1}. \quad (8)$$

In this work, all reported G_0W_0 QP energies are obtained by solving Eq. 6 self-consistently using a Newton solver. Note that in addition to the quasiparticle energies, the full G_0W_0 interacting Green's function can be obtained through Dyson's equation:

$$G^{-1}(\mathbf{r}, \mathbf{r}', \omega) = G_0^{-1}(\mathbf{r}, \mathbf{r}', \omega) - \Sigma(\mathbf{r}, \mathbf{r}', \omega), \quad (9)$$

although we do not present results for the full Green's function in this work.

B. Gaussian Density Fitting

To represent the crystalline molecular orbitals appearing in the formulae above, we employ a single-particle basis of crystalline Gaussian atomic orbitals. These are translational-symmetry-adapted linear combinations of Gaussian AOs $\{\tilde{\phi}_\mu\}$:

$$\phi_{\mu\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{T}} e^{i\mathbf{k}\cdot\mathbf{T}} \tilde{\phi}_\mu(\mathbf{r} - \mathbf{T}), \quad (10)$$

where \mathbf{T} is a lattice translation vector.

The most computationally expensive set of matrix elements in this basis to compute and store are those associated with the electron repulsion integrals (ERIs). Directly computing 4-center ERIs and using them to construct the G_0W_0 self-energy is particularly expensive in periodic systems. Therefore, we use the density fitting (resolution of identity) technique [52] to reduce the computational cost of Eqs. 3-5 and to avoid storing the full set of 4-center ERIs. We use periodic Gaussian density fitting (GDF) in the Coulomb metric as described in Ref. [48], where ERIs are decomposed into 3-center 2-electron integrals:

$$(p\mathbf{k}_p q\mathbf{k}_q | r\mathbf{k}_r s\mathbf{k}_s) = \sum_{PQ} (p\mathbf{k}_p q\mathbf{k}_q | P\mathbf{k}_{pq}) \mathbf{J}_{PQ}^{-1} (Q\mathbf{k}_{rs} | r\mathbf{k}_r s\mathbf{k}_s), \quad (11)$$

where

$$\mathbf{J}_{PQ}(\mathbf{k}) = \iint d\mathbf{r} d\mathbf{r}' \phi_{P(-\mathbf{k})}(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \phi_{Q\mathbf{k}}(\mathbf{r}'), \quad (12)$$

$$(\mathcal{Q}k_{rs}|rk_r,sk_s) = \iint d\mathbf{r}d\mathbf{r}' \phi_{\mathcal{Q}k_{rs}}(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} \phi_{rk_r}^*(\mathbf{r}') \phi_{sk_s}(\mathbf{r}'). \quad (13)$$

Here, p, q, r, s are crystalline Gaussian AOs and P, Q denote auxiliary periodic Gaussian functions. $\mathbf{k}_{pq}, \mathbf{k}_p$ and \mathbf{k}_q satisfy momentum conservation: $\mathbf{k}_{pq} = \mathbf{k}_p - \mathbf{k}_q + N\mathbf{b}$, where \mathbf{b} is a reciprocal lattice vector and N is an integer. Because of the momentum conservation relation $\mathbf{k}_p - \mathbf{k}_q + \mathbf{k}_r - \mathbf{k}_s = N\mathbf{b}$, it is also clear that $\mathbf{k}_{pq} = -\mathbf{k}_{rs}$. We note that no complex conjugate needs to be used for the auxiliary orbitals in the integrals, since the phases in the crystalline AOs are completely determined by the crystal momentum.

By decomposing the inverse Coulomb matrix $\mathbf{J}_{PQ}^{-1} = \sum_R \mathbf{J}_{PR}^{-1/2} \mathbf{J}_{RQ}^{-1/2}$, Eq. 11 is further simplified to

$$(pk_p qk_q | rk_r, sk_s) = \sum_R v_{Rk_{pq}}^{pk_p, qk_q} \cdot v_{Rk_{rs}}^{rk_r, sk_s}, \quad (14)$$

where

$$v_{Rk_{rs}}^{rk_r, sk_s} = \sum_Q \mathbf{J}_{RQ}^{-\frac{1}{2}}(\mathbf{k}_{rs}) \cdot (\mathcal{Q}k_{rs} | rk_r, sk_s). \quad (15)$$

In practice, $v_{Rk_{rs}}^{rk_r, sk_s}$ is computed and stored as $v_R^{rk_r, sk_s}$, because \mathbf{k}_{rs} is always determined by \mathbf{k}_r and \mathbf{k}_s . To maintain consistency with plane-wave expressions in the G_0W_0 literature, we can relabel the $v_R^{rk_r, sk_s}$ integral as $v_{R\mathbf{q}}^{rk, sk-\mathbf{q}}$, where \mathbf{k} and \mathbf{q} are crystal momentum vectors. The size of the auxiliary Gaussian basis is normally 3-10 times the size of the Gaussian AO basis, depending on whether we use an optimized auxiliary basis, or a brute-force even-tempered fitting set. We will describe how to utilize the GDF integrals in the G_0W_0 expressions in the next two sections.

C. Analytic Continuation

Numerical integration along the real frequency axis, as expressed in Eq. 1, is challenging because G_0 and W_0 both have many poles along the real axis. One way to avoid this problem is to perform the integration along the imaginary frequency axis [34, 36]:

$$\Sigma(\mathbf{r}, \mathbf{r}', i\omega) = -\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' G_0(\mathbf{r}, \mathbf{r}', i\omega + i\omega') W_0(\mathbf{r}, \mathbf{r}', i\omega'), \quad (16)$$

and then to analytically continue the self-energy to the real axis to computing QP energies and other spectral quantities.

The non-interacting Green's function G_0 on the imaginary axis becomes

$$G_0(\mathbf{r}, \mathbf{r}', i\omega) = \sum_{m\mathbf{k}_m} \frac{\psi_{m\mathbf{k}_m}(\mathbf{r}) \psi_{m\mathbf{k}_m}^*(\mathbf{r}')}{i\omega + \epsilon_F - \epsilon_{m\mathbf{k}_m}}. \quad (17)$$

For gapped systems, we take the Fermi energy ϵ_F to be the midpoint between the DFT valence band maximum and conduction band minimum energies.

The self-energy matrix elements in the MO basis are then computed as

$$\Sigma_{nn'}(\mathbf{k}, i\omega) = -\frac{1}{2\pi N_{\mathbf{k}}} \sum_{m\mathbf{q}} \int_{-\infty}^{\infty} d\omega' [\mathbf{G}_0(\mathbf{k} - \mathbf{q}, i\omega + i\omega')]_{mm} \times (n\mathbf{k}, m\mathbf{k} - \mathbf{q} | W_0 | m\mathbf{k} - \mathbf{q}, n'\mathbf{k}). \quad (18)$$

Here, $N_{\mathbf{k}}$ is the number of sampled \mathbf{k} -points, indices n, n' and m refer to molecular orbitals. Molecular orbitals n and n' share the same \mathbf{k} and m appears with $\mathbf{k} - \mathbf{q}$. To compute the matrix elements of W_0 in Eq. 18, we rewrite Eqs. 3 and 4 as an infinite summation by a Taylor expansion:

$$W_0 = v + v\chi_0 v + v\chi_0 v\chi_0 v + \dots, \quad (19)$$

where $v\chi_0 v$ and $v\chi_0 v\chi_0 v$ involve integrations over real-space coordinates. Meanwhile, we expand the orbital pair product $\psi_{n\mathbf{k}}^*(\mathbf{r}) \psi_{m\mathbf{k}-\mathbf{q}}(\mathbf{r})$ in the auxiliary basis

$$\psi_{n\mathbf{k}}^*(\mathbf{r}) \psi_{m\mathbf{k}-\mathbf{q}}(\mathbf{r}) = \sum_P b_{P\mathbf{q}}^{n\mathbf{k}, m\mathbf{k}-\mathbf{q}} \phi_{P\mathbf{q}}(\mathbf{r}), \quad (20)$$

with

$$b_{P\mathbf{q}}^{n\mathbf{k}, m\mathbf{k}-\mathbf{q}} = \sum_R (n\mathbf{k}, m\mathbf{k} - \mathbf{q} | R\mathbf{q}) \cdot \mathbf{J}_{RP}^{-1}(\mathbf{q}). \quad (21)$$

To ease notation, some momentum labels are suppressed in the above and following equations (e.g., we will use b_P^{nm} to denote $b_{P\mathbf{q}}^{n\mathbf{k}, m\mathbf{k}-\mathbf{q}}$). Using Eqs. 19-21, the matrix elements of W_0 are computed as

$$\begin{aligned} & (n\mathbf{k}, m\mathbf{k} - \mathbf{q} | W_0 | m\mathbf{k} - \mathbf{q}, n'\mathbf{k}) \\ &= \sum_{PQ} b_P^{nm} \left[\iint d\mathbf{r}d\mathbf{r}' \phi_{P\mathbf{q}}(\mathbf{r}) W_0(\mathbf{r}, \mathbf{r}', i\omega') \phi_{Q(-\mathbf{q})}(\mathbf{r}') \right] b_Q^{m\mathbf{k}-\mathbf{q}, n'\mathbf{k}} \\ &= \sum_{PQ} b_P^{nm} [\mathbf{J}_{PQ}(\mathbf{q}) + (\mathbf{J}^{1/2} \mathbf{\Pi} \mathbf{J}^{1/2})_{PQ}(\mathbf{q}) + \dots] b_Q^{m\mathbf{k}-\mathbf{q}, n'\mathbf{k}} \\ &= \sum_{PQ} v_P^{nm} [\mathbf{I} - \mathbf{\Pi}(\mathbf{q}, i\omega')]_{PQ}^{-1} v_Q^{m\mathbf{k}-\mathbf{q}, n'\mathbf{k}}. \end{aligned} \quad (22)$$

The 3-center 2-electron integral v_P^{nm} between auxiliary basis function P and molecular orbital pairs nm is obtained from an AO to MO transformation of the GDF AO integrals defined in Eq. 15:

$$v_P^{nm} = \sum_p \sum_q C_{pn}(\mathbf{k}) C_{qm}(\mathbf{k} - \mathbf{q}) v_{P\mathbf{q}}^{pk, qk-\mathbf{q}}, \quad (23)$$

where $C(\mathbf{k})$ refers to the MO coefficients in the AO basis. $\mathbf{\Pi}(\mathbf{q}, i\omega')$ in Eq. 22 is an auxiliary density response function:

$$\mathbf{\Pi}_{PQ}(\mathbf{q}, i\omega') = \frac{2}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_i^{\text{occ}} \sum_a^{\text{vir}} v_P^{ia} \frac{\epsilon_{i\mathbf{k}} - \epsilon_{a\mathbf{k}-\mathbf{q}}}{\omega'^2 + (\epsilon_{i\mathbf{k}} - \epsilon_{a\mathbf{k}-\mathbf{q}})^2} v_Q^{ai}. \quad (24)$$

Inserting Eq. 17 and Eq. 22, Eq. 18 becomes

$$\begin{aligned} \Sigma_{mn'}(\mathbf{k}, i\omega) = & -\frac{1}{2\pi N_{\mathbf{k}}} \sum_{m\mathbf{q}} \int_{-\infty}^{\infty} d\omega' \frac{1}{i(\omega + \omega') + \epsilon_F - \epsilon_{m\mathbf{k}-\mathbf{q}}} \\ & \times \sum_{PQ} v_P^{nm} [\mathbf{I} - \mathbf{\Pi}(\mathbf{q}, i\omega')]_{PQ}^{-1} v_Q^{mn'}. \end{aligned} \quad (25)$$

The self-energy term in Eq. 25 is further divided into exchange and correlation components $\Sigma(\mathbf{k}, i\omega) = \Sigma^x(\mathbf{k}) + \Sigma^c(\mathbf{k}, i\omega)$, where the frequency-independent exchange $\Sigma^x(\mathbf{k})$ is the Hartree-Fock (HF) exchange matrix evaluated using the DFT orbitals:

$$\Sigma_{mn'}^x(\mathbf{k}) = -\frac{1}{N_{\mathbf{k}}} \sum_{P\mathbf{q}} \sum_i^{\text{occ}} v_{P\mathbf{q}}^{nk, ik-\mathbf{q}} \cdot v_{P(-\mathbf{q})}^{ik-\mathbf{q}, n'\mathbf{k}}. \quad (26)$$

The advantage of this division is that the HF exchange is free of integration error. Accordingly, the correlation part of self-energy becomes

$$\begin{aligned} \Sigma_{mn'}^c(\mathbf{k}, i\omega) = & -\frac{1}{\pi N_{\mathbf{k}}} \sum_{m\mathbf{q}} \int_0^{\infty} d\omega' \frac{i\omega + \epsilon_F - \epsilon_{m\mathbf{k}-\mathbf{q}}}{(i\omega + \epsilon_F - \epsilon_{m\mathbf{k}-\mathbf{q}})^2 + \omega'^2} \\ & \times \sum_{PQ} v_P^{nm} [\mathbf{I} - \mathbf{\Pi}(\mathbf{q}, i\omega')]_{PQ}^{-1} - \delta_{PQ} v_Q^{mn'}. \end{aligned} \quad (27)$$

The integration in Eq. 27 must be performed on a numerical grid. Following Ref. [36], we employ a modified Gauss-Legendre grid that transforms a standard Gauss-Legendre grid in the range $[-1, 1]$ to $[0, \infty]$:

$$\tilde{x}_i = \frac{1 + x_i}{2(1 - x_i)}, \quad \tilde{w}_i = \frac{w_i}{(1 - x_i)^2}. \quad (28)$$

$\{x_i, w_i\}$ are the original Gauss-Legendre abscissas and weights of the grid points in the range $[-1, 1]$, and $\{\tilde{x}_i, \tilde{w}_i\}$ are the transformed abscissas and weights of the grid points used for integrating Eq. 27. In this work, $N_G = 100$ grid points were used for all reported results unless specified.

The real-frequency self-energy is then obtained through an analytic continuation scheme by fitting to an analytic function defined on the full complex plane. We fit the self-energy matrix elements to N -point Padé approximants using Thiele's reciprocal difference method (see Appendix of Ref. [53] for the detailed algorithm):

$$\Sigma_{mn'}(\mathbf{k}, z) = \frac{a_0 + a_1 \cdot z + \dots + a_{(N-1)/2} \cdot z^{(N-1)/2}}{1 + b_1 \cdot z + \dots + b_{N/2} \cdot z^{N/2}}, \quad (29)$$

where z is any complex frequency, $\{a_i, b_i\}$ are Padé coefficients that need to be fitted. In this work, we use $N = 18$ Padé approximants, which requires 18 $\Sigma_{mn'}(\mathbf{k}, z = i\omega)$ data points as input to the fit. We choose these 18 imaginary frequency points to be the positions of the modified Gauss-Legendre grid points. Once the Padé polynomials are fitted, the real-frequency self-energy is easily calculated:

$$\Sigma_{mn'}(\mathbf{k}, \omega) = \frac{a_0 + a_1 \cdot \omega + \dots + a_{(N-1)/2} \cdot \omega^{(N-1)/2}}{1 + b_1 \cdot \omega + \dots + b_{N/2} \cdot \omega^{N/2}}. \quad (30)$$

We further comment on the computational scaling of this approach. If only G_0W_0 QP energies are required, one only needs to compute the diagonal self-energy matrix elements. The most expensive step is then computing the auxiliary density response function in Eq. 24, whose cost scales as $\mathcal{O}(N_{\mathbf{k}}^2 N_o N_v N_{aux}^2)$, where N_o, N_v, N_{aux} are the number of occupied, virtual and auxiliary orbitals per unit cell. There is also a prefactor N_G (number of grid points) in Eq. 24, which is fixed regardless of the system size. The second expensive step is integrating Eq. 27, which scales as $\mathcal{O}(N_{\mathbf{k}}^2 N_{AO} N_{aux}^2)$ and has a prefactor $N_G N_{Pade}$, with N_{AO} as the number of atomic orbitals per unit cell and N_{Pade} as the number of Padé approximants. On the other hand, if the full G_0W_0 Green's function and off-diagonal self-energy matrix at all \mathbf{k} -points are also required, Eq. 27 becomes the most time-consuming step, with a cost scaling of $\mathcal{O}(N_{\mathbf{k}}^2 N_{AO}^2 N_{aux}^2)$.

D. Contour Deformation

The AC scheme has been shown to give accurate G_0W_0 valence state energies in molecules and solids [21, 34, 54]. However, analytic continuation is known to be very unstable for states far away from the Fermi level (e.g., core excitation energies) and thus gives inaccurate G_0W_0 QP energies for those states [55]. A more robust scheme is the contour deformation (CD) approach [13, 37], where the integral over the real frequency axis in Eq. 1 is transformed into an integral over the contours shown in Fig. 1.

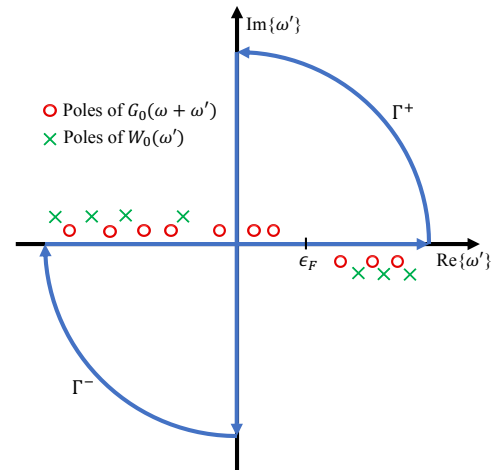


FIG. 1. Contours used for integration in the G_0W_0 -CD approach. The integration contours Γ^+ and Γ^- enclose some poles of the Green's function G_0 , but exclude poles of the screened Coulomb interaction W_0 .

The integration in Eq. 1 can then be broken down into the sum of an integration along the imaginary frequency axis plus

all residues arising from poles enclosed within the contours:

$$\begin{aligned}\Sigma(\mathbf{r}, \mathbf{r}', \omega) &= \frac{i}{2\pi} \oint d\omega' G_0(\mathbf{r}, \mathbf{r}', \omega + \omega') W_0(\mathbf{r}, \mathbf{r}', \omega') \\ &\quad - \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' G_0(\mathbf{r}, \mathbf{r}', \omega + i\omega') W_0(\mathbf{r}, \mathbf{r}', i\omega') \\ &= \Sigma^C(\mathbf{r}, \mathbf{r}', \omega) + \Sigma^I(\mathbf{r}, \mathbf{r}', \omega).\end{aligned}\quad (31)$$

From Eq. 2, the poles of $G_0(\mathbf{r}, \mathbf{r}', \omega + \omega')$ are located at frequencies

$$\omega'_{m\mathbf{k}_m} = \epsilon_{m\mathbf{k}_m} - \omega + i\eta \text{sgn}(\epsilon_F - \epsilon_{m\mathbf{k}_m}), \quad (32)$$

with residues

$$\text{Res}\{G_0(\mathbf{r}, \mathbf{r}', \omega + \omega'), \omega'_{m\mathbf{k}_m}\} = \psi_{m\mathbf{k}_m}(\mathbf{r}) \psi_{m\mathbf{k}_m}^*(\mathbf{r}'). \quad (33)$$

The poles of $G_0(\mathbf{r}, \mathbf{r}', \omega + \omega')$ thus enter the contour Γ^+ when $\epsilon_{m\mathbf{k}_m} < \epsilon_F$ and $\omega'_{m\mathbf{k}_m} > 0$. Similarly, when $\epsilon_{m\mathbf{k}_m} > \epsilon_F$ and $\omega'_{m\mathbf{k}_m} < 0$, the poles of $G_0(\mathbf{r}, \mathbf{r}', \omega + \omega')$ enter the contour Γ^- . On the other hand, the poles of $W_0(\mathbf{r}, \mathbf{r}', \omega')$ are never enclosed in the contours in Fig. 1. Therefore, the integral of the contour in Eq. 31 is computed as

$$\begin{aligned}\Sigma_{nm'}^C(\mathbf{k}, \omega) &= \frac{1}{N_{\mathbf{k}}} \sum_{m\mathbf{q}} f_{m\mathbf{k}-\mathbf{q}} \\ &\quad \times \left(n\mathbf{k}, m\mathbf{k} - \mathbf{q} \left| W(\omega'_{m\mathbf{k}-\mathbf{q}}) \right| m\mathbf{k} - \mathbf{q}, n'\mathbf{k} \right) \\ &= \frac{1}{N_{\mathbf{k}}} \sum_{m\mathbf{q}} f_{m\mathbf{k}-\mathbf{q}} \sum_{PQ} v_P^{nm} [\mathbf{I} - \mathbf{\Pi}(\mathbf{q}, \omega'_{m\mathbf{k}-\mathbf{q}})]_{PQ}^{-1} v_Q^{mn'},\end{aligned}\quad (34)$$

and the contribution of residues $f_{m\mathbf{k}-\mathbf{q}}$ is given by

$$f_{m\mathbf{k}-\mathbf{q}} = \begin{cases} 1 & \text{if } \epsilon_F < \epsilon_{m\mathbf{k}-\mathbf{q}} < \omega; \\ -1 & \text{if } \epsilon_F > \epsilon_{m\mathbf{k}-\mathbf{q}} > \omega; \\ 0 & \text{else.} \end{cases} \quad (35)$$

The auxiliary density response function is computed as (we use $\eta = 0.001$ a.u.)

$$\begin{aligned}\Pi_{PQ}(\mathbf{q}, \omega) &= \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \sum_i^{\text{occ}} \sum_a^{\text{vir}} v_P^{ia} \left(\frac{1}{\omega - (\epsilon_{a\mathbf{k}-\mathbf{q}} - \epsilon_{i\mathbf{k}}) + i\eta} \right. \\ &\quad \left. - \frac{1}{\omega + (\epsilon_{a\mathbf{k}-\mathbf{q}} - \epsilon_{i\mathbf{k}}) - i\eta} \right) v_Q^{ai}.\end{aligned}\quad (36)$$

The other integration over the imaginary frequency axis in Eq. 31 is calculated on the modified Gauss-Legendre grid, as described in Section II C:

$$\begin{aligned}\Sigma_{nm'}^I(\mathbf{k}, \omega) &= -\frac{1}{2\pi N_{\mathbf{k}}} \sum_{m\mathbf{k}-\mathbf{q}} \int_{-\infty}^{\infty} d\omega' \frac{1}{\omega + i\omega' + \epsilon_F - \epsilon_{m\mathbf{k}-\mathbf{q}}} \\ &\quad \times \sum_{PQ} v_P^{nm} [\mathbf{I} - \mathbf{\Pi}(\mathbf{q}, i\omega')]_{PQ}^{-1} v_Q^{mn'},\end{aligned}\quad (37)$$

where the expression of $\mathbf{\Pi}(\mathbf{q}, i\omega')$ follows Eq. 24. We note that similar to Eqs. 26 and 27, we also separately compute the exchange and correlation parts of self-energy in Eqs. 34 and 37.

The G_0W_0 -CD approach has the advantage that the real-axis self-energy is directly computed without the need for analytic continuation. However, the computational expense is higher than in the G_0W_0 -AC approach. It is clear that computing the imaginary integration alone in Eq. 37 has a similar computational cost to the G_0W_0 -AC scheme. The extra cost comes from computing the pole residues in Eq. 34, which become more expensive if more poles need to be calculated (i.e., if the targeted states are further away from the Fermi level). An extreme case is the core excitations, where the computational scaling of Eq. 34 is approximately $\mathcal{O}(N_{\mathbf{k}}^3 N_o^2 N_v N_{aux}^2)$. Nevertheless, as shown in previous literature, the G_0W_0 -CD scheme needs to be chosen over the G_0W_0 -AC scheme in order to obtain numerically stable core excitation energies in molecules [37, 56].

E. Coulomb Divergence Correction

The periodic GDF Coulomb integral $v_{R\mathbf{q}}^{r\mathbf{k}, s\mathbf{k}-\mathbf{q}}$ removes the $\mathbf{G} = \mathbf{0}$ contribution (here \mathbf{G} is a reciprocal lattice vector) when $\mathbf{q} = \mathbf{0}$, because the Coulomb interaction $4\pi/\Omega_{\mathbf{G}^2}$ diverges at $\mathbf{G} = \mathbf{0}$. This leads to an $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ finite size error in our G_0W_0 calculations. This component of the finite size error in the exchange self-energy matrix can be corrected by shifting the occupied orbital energies entering into G_0 using the probe-charge Ewald method [57] (exxdiv='ewald' in PySCF). To correct the remaining leading order finite size error in the correlation self-energy, we rewrite the G_0W_0 -AC self-energy expression in Eq. 27 in a plane-wave expansion [23]:

$$\begin{aligned}\Sigma_{nm'}^C(\mathbf{k}, i\omega) &= -\frac{1}{2\pi N_{\mathbf{k}} \Omega_{\text{cell}}} \sum_{m\mathbf{q}} \sum_{\mathbf{G}\mathbf{G}'} \int_{-\infty}^{\infty} d\omega' W_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega') \\ &\quad \times \frac{\rho_{m\mathbf{k}-\mathbf{q}}^{n\mathbf{k}}(\mathbf{G}) \rho_{m\mathbf{k}-\mathbf{q}}^{n'\mathbf{k}*}(\mathbf{G}')}{i(\omega + \omega') + \epsilon_F - \epsilon_{m\mathbf{k}-\mathbf{q}}},\end{aligned}\quad (38)$$

where Ω_{cell} is the unit cell volume. The pair density nm is now expanded in plane waves (\mathbf{G} and \mathbf{G}'), with the matrix elements

$$\rho_{m\mathbf{k}-\mathbf{q}}^{n\mathbf{k}}(\mathbf{G}) = (n\mathbf{k} | e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}} | m\mathbf{k} - \mathbf{q}). \quad (39)$$

The screened Coulomb potential is computed as:

$$W_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega) = \frac{\sqrt{4\pi}}{|\mathbf{q} + \mathbf{G}|} (\epsilon_{\mathbf{G}\mathbf{G}'}^{-1}(\mathbf{q}, i\omega) - \delta_{\mathbf{G}\mathbf{G}'}) \frac{\sqrt{4\pi}}{|\mathbf{q} + \mathbf{G}'|}, \quad (40)$$

with the dielectric function defined as:

$$\epsilon_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega) = \delta_{\mathbf{G}\mathbf{G}'} - \frac{\sqrt{4\pi}}{|\mathbf{q} + \mathbf{G}|} \chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega) \frac{\sqrt{4\pi}}{|\mathbf{q} + \mathbf{G}'|}, \quad (41)$$

and the polarizability kernel is computed as:

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega) = \frac{2}{N_{\mathbf{k}}\Omega_{\text{cell}}} \sum_{\mathbf{k}} \sum_i^{\text{occ}} \sum_a^{\text{vir}} \rho_{\mathbf{ak}-\mathbf{q}}^{ik}(\mathbf{G}) \quad (42)$$

$$\times \frac{\epsilon_{ik} - \epsilon_{\mathbf{ak}-\mathbf{q}}}{\omega^2 + (\epsilon_{ik} - \epsilon_{\mathbf{ak}-\mathbf{q}})^2} \rho_{\mathbf{ak}-\mathbf{q}}^{ik*}(\mathbf{G}').$$

In Eq. 42, when $\mathbf{G} = \mathbf{0}$, $\rho_{\mathbf{ak}-\mathbf{q}}^{ik}(\mathbf{G}) \approx i\mathbf{q} \cdot (i\mathbf{k}|\mathbf{r}|\mathbf{ak}-\mathbf{q})$ at $\mathbf{q} \rightarrow \mathbf{0}$, so $\chi_{\mathbf{0}\mathbf{0}}(\mathbf{q} \rightarrow \mathbf{0}, i\omega) = \mathcal{O}(\mathbf{q}^2)$. Therefore, in Eq. 41, $\epsilon_{\mathbf{0}\mathbf{0}}(\mathbf{q} \rightarrow \mathbf{0}, i\omega)$ has finite value. This means the head of the screened Coulomb potential $W_{\mathbf{0}\mathbf{0}}$ diverges as $\mathcal{O}(1/\mathbf{q}^2)$ when $\mathbf{q} \rightarrow \mathbf{0}$. Similarly, the wings of the screened Coulomb potential $W_{\mathbf{G}\mathbf{0}}$ and $W_{\mathbf{0}\mathbf{G}'}$ diverge as $\mathcal{O}(1/\mathbf{q})$. However, in the limit of a very fine \mathbf{k} -point sampling, $\sum_{\mathbf{q}} \rightarrow \frac{\Omega}{(2\pi)^3} \int dq 4\pi q^2$ ($\Omega = \Omega_{\text{cell}} N_{\mathbf{k}}$), thus $W_{\mathbf{0}\mathbf{0}}$ and $W_{\mathbf{G}\mathbf{0}}$ are integrable.

Following Refs. [17, 23, 41], we add head and wings finite size corrections in our G_0W_0 implementation. We determine the contributions for $\mathbf{q} = \mathbf{0}$, $\mathbf{G} = \mathbf{0}$ by analytically integrating the $\mathbf{q} \rightarrow \mathbf{0}$ contributions in a sphere of radius $\mathbf{q} = q_0$ around $\mathbf{G} = \mathbf{0}$ ($\frac{\Omega}{(2\pi)^3} \int_0^{q_0} dq 4\pi q^2 = 1$). Doing this for the head of screened Coulomb potential in Eq. 40 gives

$$W_{\mathbf{0}\mathbf{0}}(\mathbf{q} = \mathbf{0}, i\omega) = \frac{2\Omega}{\pi} \left(\frac{6\pi^2}{\Omega} \right)^{1/3} [\epsilon_{\mathbf{0}\mathbf{0}}^{-1}(\mathbf{q} = \mathbf{0}, i\omega) - 1]. \quad (43)$$

In the $\mathbf{q} \rightarrow \mathbf{0}$ limit, $\rho_{m\mathbf{k}-\mathbf{q}}^{nk}(\mathbf{0}) \approx 1$ only at $n = m$. Therefore, inserting Eq. 43 into Eq. 38, one arrives at the head correction to the self-energy (which only modifies the diagonal self-energy matrix elements):

$$\Sigma_{nn}^{\text{head}}(\mathbf{k}, i\omega) = -\frac{1}{\pi^2} \left(\frac{6\pi^2}{\Omega_{\text{cell}} N_{\mathbf{k}}} \right)^{1/3} \quad (44)$$

$$\times \int_{-\infty}^{\infty} d\omega' \frac{\epsilon_{\mathbf{0}\mathbf{0}}^{-1}(\mathbf{q} = \mathbf{0}, i\omega) - 1}{i(\omega + \omega') + \epsilon_F - \epsilon_{n\mathbf{k}}},$$

where we assume $\epsilon_{n\mathbf{k}-\mathbf{q}} = \epsilon_{n\mathbf{k}}$. From Eq. 44 one can clearly see that ignoring this term leads to an $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$ finite size error. Similarly, we can obtain the wings contribution to the self-energy:

$$\Sigma_{nn'}^{\text{wings}}(\mathbf{k}, i\omega) = -\frac{1}{\pi} \sqrt{\frac{\Omega_{\text{cell}}}{4\pi^3}} \left(\frac{6\pi^2}{\Omega_{\text{cell}} N_{\mathbf{k}}} \right)^{2/3} \quad (45)$$

$$\times \sum_P \int_{-\infty}^{\infty} d\omega' \frac{\text{Re}[v_P^{nn'} \epsilon_{P\mathbf{0}}^{-1}(\mathbf{q} = \mathbf{0}, i\omega)]}{i(\omega + \omega') + \epsilon_F - \epsilon_{n'\mathbf{k}}}.$$

Here, P refers to the Gaussian auxiliary basis and $v_P^{nn'}$ is the GDF integral defined as $v_{P\mathbf{0}}^{n\mathbf{k}, n'\mathbf{k}}$.

The head and wings of the dielectric function in the $\mathbf{q} = \mathbf{0}$ limit are computed as

$$h_{\mathbf{0}\mathbf{0}}(\mathbf{q} = \mathbf{0}, i\omega) = 1 - \lim_{\mathbf{q} \rightarrow \mathbf{0}} \frac{4\pi}{|\mathbf{q}|^2} \chi_{\mathbf{0}\mathbf{0}}(\mathbf{q} \rightarrow \mathbf{0}, i\omega), \quad (46)$$

$$w_{P\mathbf{0}}(\mathbf{q} = \mathbf{0}, i\omega) = \lim_{\mathbf{q} \rightarrow \mathbf{0}} \frac{\sqrt{4\pi}}{|\mathbf{q}|} \chi_{P\mathbf{0}}(\mathbf{q} \rightarrow \mathbf{0}, i\omega), \quad (47)$$

where the polarizability is

$$\chi_{\mathbf{0}\mathbf{0}}(\mathbf{q} \rightarrow \mathbf{0}, i\omega) = \frac{2}{N_{\mathbf{k}}\Omega_{\text{cell}}} \sum_{\mathbf{k}} \sum_i^{\text{occ}} \sum_a^{\text{vir}} \rho_{\mathbf{ak}-\mathbf{q}}^{ik}(\mathbf{0}) \quad (48)$$

$$\times \frac{\epsilon_{ik} - \epsilon_{\mathbf{ak}}}{\omega^2 + (\epsilon_{ik} - \epsilon_{\mathbf{ak}})^2} \rho_{\mathbf{ak}-\mathbf{q}}^{ik*}(\mathbf{0}),$$

$$\chi_{P\mathbf{0}}(\mathbf{q} \rightarrow \mathbf{0}, i\omega) = \frac{2}{N_{\mathbf{k}}\Omega_{\text{cell}}^{1/2}} \sum_{\mathbf{k}} \sum_i^{\text{occ}} \sum_a^{\text{vir}} v_P^{ia} \quad (49)$$

$$\times \frac{\epsilon_{ik} - \epsilon_{\mathbf{ak}}}{\omega^2 + (\epsilon_{ik} - \epsilon_{\mathbf{ak}})^2} \rho_{\mathbf{ak}-\mathbf{q}}^{ik*}(\mathbf{0}).$$

In Eqs. 46 and 47 we evaluate the limit using $\mathbf{q} = (0.001, 0, 0)$; we neglect the anisotropy of this limit for simplicity in this work, although anisotropic corrections can be obtained as discussed in Ref. [58]. The pair density matrix in the long-wavelength limit is computed using $\mathbf{k} \cdot \mathbf{p}$ perturbation theory as described in Ref. [59]:

$$\rho_{\mathbf{ak}-\mathbf{q}}^{ik}(\mathbf{0}) \Big|_{\mathbf{q} \rightarrow \mathbf{0}} = \frac{-i\mathbf{q} \cdot (\psi_{ik} | \nabla | \psi_{\mathbf{ak}})}{\epsilon_{\mathbf{ak}} - \epsilon_{ik}}. \quad (50)$$

Finally, we note that $\epsilon_{\mathbf{0}\mathbf{0}}^{-1}(\mathbf{q} = \mathbf{0}, i\omega)$ and $\epsilon_{P\mathbf{0}}^{-1}(\mathbf{q} = \mathbf{0}, i\omega)$ in Eqs. 44 and 45 are not simple inverses of $h_{\mathbf{0}\mathbf{0}}$ and $w_{P\mathbf{0}}$ in Eqs. 46 and 47. Instead, they are matrix elements of the inverse of the full dielectric matrix [17, 41]:

$$\epsilon_{\text{full}}(\mathbf{q} = \mathbf{0}, i\omega) = \begin{bmatrix} h_{\mathbf{0}\mathbf{0}} & w_{P\mathbf{0}}^\dagger \\ w_{P\mathbf{0}} & B_{PQ} \end{bmatrix}, \quad (51)$$

where $B_{PQ} = [\mathbf{I} - \mathbf{\Pi}]_{PQ}$ is the body of the dielectric function computed in the Gaussian auxiliary basis (see Eq. 24). By inverting Eq. 51, one obtains $\epsilon_{\mathbf{0}\mathbf{0}}^{-1}(\mathbf{q} = \mathbf{0}, i\omega)$ and $\epsilon_{P\mathbf{0}}^{-1}(\mathbf{q} = \mathbf{0}, i\omega)$ as

$$\epsilon_{\mathbf{0}\mathbf{0}}^{-1} = 1 / (h_{\mathbf{0}\mathbf{0}} - \sum_{PQ} w_{P\mathbf{0}}^\dagger B_{PQ}^{-1} w_{Q\mathbf{0}}), \quad (52)$$

$$\epsilon_{P\mathbf{0}}^{-1} = -\epsilon_{\mathbf{0}\mathbf{0}}^{-1} \sum_Q B_{PQ}^{-1} w_{Q\mathbf{0}}. \quad (53)$$

The finite size corrections in the G_0W_0 -CD approach are also implemented in a similar manner.

III. G_0W_0 BENCHMARK RESULTS

We benchmark our Gaussian-based G_0W_0 method, as implemented in PySCF, on both molecules and periodic crystals. We use standard all-electron Gaussian basis sets (all Gaussian basis sets can be found in the Basis Set Exchange database [60]), and we include all electrons in the G_0W_0 calculations. We note that our implementation can also be used in Gaussian bases with pseudopotentials [61] or effective core potentials [62], although we do not present such calculations in this work.

A. Validation of G_0W_0 -AC Code

We first validate our G_0W_0 -AC code for molecules (i.e. without periodic boundary conditions) drawn from the $GW100$ test set [54]. Ionization potentials (IP) and QP lowest molecular orbital energies (LUMO) were calculated for 15 molecules from the test set. Following Ref. [54], we used the def2-QZVP basis [63] as the orbital basis and the def2-QZVP-RIFIT basis [64] as the auxiliary basis, and the PBE density functional [65] was chosen to provide the starting choice of DFT orbitals and energies. Our G_0W_0 @PBE results are shown in Table I, and compared to those results listed as AIMS-P16 and TM-RI in Ref. [54] (note the TM-RI implementation also used Gaussian density fitting ERIs). As shown in the table, our G_0W_0 @PBE results are in very good agreement with the reference data, confirming the accuracy of our implementation for molecules.

TABLE I. G_0W_0 @PBE IP and QP-LUMO energies of 15 molecules from the $GW100$ set [54] in the def2-QZVP basis using the G_0W_0 -AC approach. All numbers are in eV. MAD stands for mean absolute difference between two methods.

Molecule	IP	QP-LUMO
CH ₄	13.93	2.45
H ₂ O	11.97	2.37
SiH ₄	12.31	2.51
LiH	6.52	-0.07
CO	13.57	0.67
CO ₂	13.25	2.50
SO ₂	11.82	-1.00
N ₂	14.89	2.45
P ₂	10.21	-0.72
Cu ₂	7.52	-0.96
NaCl	8.10	-0.39
BrK	7.33	-0.31
TiF ₄	13.90	-0.60
C ₆ H ₆	8.99	1.09
C ₅ H ₅ N ₅ O (guanine)	7.69	0.75
MAD (PySCF – AIMS-P16 [54])	0.01	0.00
MAD (PySCF – TM-RI [54])	0.04	0.03

B. AC vs. CD Schemes

We next apply our periodic G_0W_0 -AC and G_0W_0 -CD implementations to 11 prototypical semiconductors (Si and C in the diamond structure; SiC, AlP, BN, BP, GaN, GaP, ZnO and ZnS in the zinc blende structure; MgO in the rock salt structure) and 2 rare gas solids (Ne and Ar in the fcc structure). The aim here is to compare the accuracy of G_0W_0 -AC against the more robust G_0W_0 -CD method for obtaining valence state energies, rather than obtaining converged band gap results (which will be presented in Section III E). Thus, here we used Dunning’s correlation consistent cc-pVTZ basis [66–70] (cc-pVTZ-RI [64, 71, 72] was used as the auxiliary basis) and a moderate $4 \times 4 \times 4$ \mathbf{k} -mesh for all the tested systems. For MgO, the cc-pVTZ basis exhibits severe linear dependencies, so cc-pVDZ/cc-pVDZ-RI basis sets were used instead. For the

rare gas solids, aug-cc-pVTZ/aug-cc-pVTZ-RI basis sets [73] were used to accurately treat the van der Waals interactions. No finite size corrections for the exchange and correlation self-energies were applied. The G_0W_0 @PBE valence band maximum (VBM) and conduction band minimum (CBM) energies of tested solids from the two schemes are presented in Table II.

TABLE II. G_0W_0 @PBE valence band maximum and conduction band minimum energies for several periodic crystals using a $4 \times 4 \times 4$ \mathbf{k} -mesh (with no finite size corrections). The “VBM-AC” and “CBM-AC” columns are QP energies from the G_0W_0 -AC scheme, while the “ δ VBM” and “ δ CBM” columns refer to differences between the G_0W_0 -AC and G_0W_0 -CD QP energies ($\epsilon_{AC} - \epsilon_{CD}$). The unit of energy is eV.

System	VBM-AC	δ VBM ($\times 10^{-3}$)	CBM-AC	δ CBM ($\times 10^{-3}$)
Si	8.56	0	9.51	0
C	15.09	0	20.11	-1
SiC	12.08	-1	14.09	-20
AlP	7.25	0	9.42	-8
BN	12.65	0	18.27	-37
BP	11.49	-1	13.39	-14
GaN	12.23	0	15.13	-14
GaP	9.58	0	11.68	-7
MgO	8.71	0	15.33	-11
ZnO	9.08	0	11.78	0
ZnS	7.76	0	11.20	-5
Ne	-12.50	0	5.20	0
Ar	-4.86	1	6.97	6
MAD		0		10

We find that for the VBM energies, the G_0W_0 -AC results are in excellent agreement with the G_0W_0 -CD values. For the CBM energies, there is a mean absolute difference of only 0.01 eV, and the maximum difference is less than 0.04 eV, which is much smaller than the possible errors introduced from other factors (e.g., the finite basis and \mathbf{k} -mesh). Therefore, we will use the more efficient G_0W_0 -AC approach for all studies of valence band energies in the subsequent sections.

C. Basis Set Convergence

We now study how the G_0W_0 band gaps of periodic systems converge with respect to the Gaussian basis set size. This problem is especially interesting in transition metal containing systems because such systems normally require a large number of plane waves (and empty states) to converge the G_0W_0 QP energies [28, 74]. Here, we performed G_0W_0 @PBE calculations for the band gaps of Si, C and ZnO (zinc blende) using a $4 \times 4 \times 4$ \mathbf{k} -mesh. The cc-pVXZ basis sets (X=D, T, Q) and their corresponding cc-pVXZ-RI auxiliary basis were used because they are designed to capture correlation in a systematic fashion in wavefunction based quantum chemistry calculations. The results are shown in Fig. 2.

It can be seen that the G_0W_0 @PBE band gaps of Si and C are already converged to within 0.04 eV of the largest basis result using the cc-pVDZ basis, which only has 36 (Si) and 28 (C) Gaussian basis functions per unit cell. For cubic ZnO, the

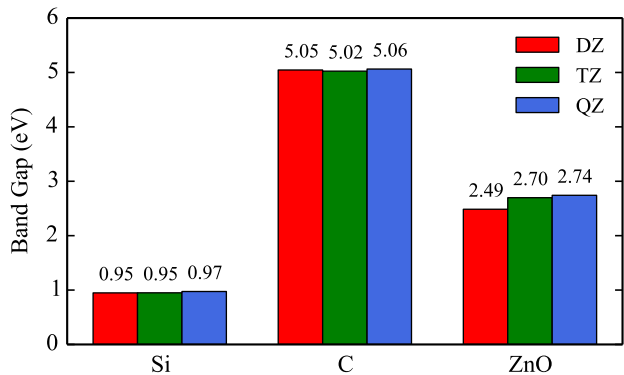


FIG. 2. G_0W_0 @PBE band gaps of Si, C and cubic ZnO using the cc-pVXZ (X=D, T, Q) basis sets and a $4 \times 4 \times 4$ \mathbf{k} -mesh (no finite size correction is applied). The band gap values are shown on top of the bars.

cc-pVTZ basis is needed to reach an accurate band gap compared to the cc-pVQZ value. However, we note that ZnO is a well-known challenging system for GW calculations, largely because the GW self-energy converges very slowly with respect to the truncated number of virtual bands in the polarizability calculation [26, 27, 74, 75]. Typically, a few thousand plane-wave basis functions are required to avoid a substantial underestimation of the G_0W_0 band gap for ZnO. One approach to deal with this issue is to extend the LAPW method with high-energy local orbitals (HLOs), which reduces the number of plane-waves to a few hundred [76]. In our Gaussian-based G_0W_0 calculation for cubic ZnO, there are only 98 (cc-pVTZ) and 159 (cc-pVQZ) basis functions per unit cell, which is much smaller than the basis size in purely plane-wave based calculations. This is because the correlation consistent construction of the Gaussian bases discretizes the virtual states in a way that is specifically designed to rapidly converge the correlation energy. Due to this small basis size, no virtual band truncation is needed to compute the polarizability. This suggests that Gaussian bases have an advantage as a compact and efficient choice when performing G_0W_0 calculations on periodic systems. From the above analysis, we will use the cc-pVTZ/cc-pVTZ-RI basis in the calculations that follow unless otherwise specified.

D. Finite Size Convergence

We next investigate the convergence of G_0W_0 valence excitation energies and band gaps with respect to the Brillouin zone sampling, i.e., the number of \mathbf{k} -points $N_{\mathbf{k}}$. As discussed in Section II E, without the $\mathbf{G} = \mathbf{0}$ contribution to the exchange and correlation G_0W_0 self-energies, we expect the G_0W_0 excitation energies and band gaps to have a finite-size error that scales like $\mathcal{O}(N_{\mathbf{k}}^{-1/3})$. We thus plot the VBM, CBM and band gaps of Si and cubic BN computed using G_0W_0 @PBE in Fig. 3 as a function of $N_{\mathbf{k}}^{-1/3}$, using the cc-pVTZ basis and increasing \mathbf{k} -meshes from $2 \times 2 \times 2$ to $7 \times 7 \times 7$. As a comparison, we also show the finite-size-corrected G_0W_0 energies obtained

by using the head and wings corrections to the correlation self-energy as well as the probe-charge Ewald correction to the exchange self-energy.

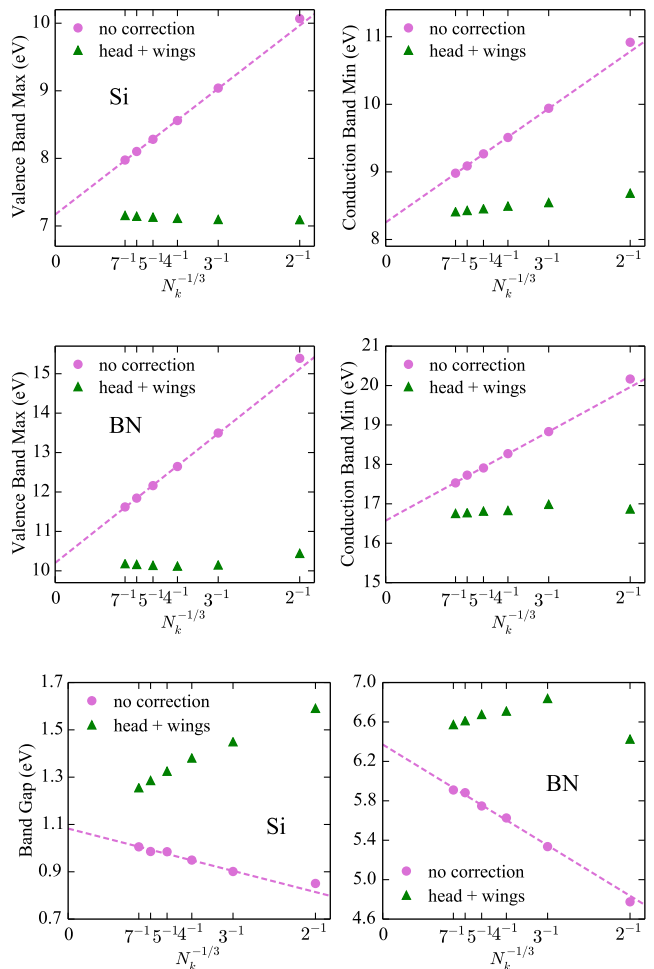


FIG. 3. G_0W_0 @PBE VBM, CBM and band gaps of Si and cubic BN as a function of $N_{\mathbf{k}}^{-1/3}$ in the cc-pVTZ basis, with and without finite size corrections. A linear extrapolation (dashed line) is performed on the uncorrected energies. Top: VBM and CBM of Si. Middle: VBM and CBM of cubic BN. Bottom: Band gaps of Si and BN.

From this data we see that the uncorrected G_0W_0 VBM, CBM and band gaps of Si and cubic BN indeed depend approximately linearly on $N_{\mathbf{k}}^{-1/3}$. We thus fit the uncorrected G_0W_0 VBM, CBM and band gap values from the $3 \times 3 \times 3$ to $7 \times 7 \times 7$ \mathbf{k} -meshes to the form $E(N_{\mathbf{k}}^{-1/3}) = E_{\infty} + aN_{\mathbf{k}}^{-1/3}$ to extrapolate to the thermodynamic limit (TDL). Compared to the extrapolated values, one can see that the finite size errors in the VBM and CBM of Si and cubic BN are significantly reduced by applying the finite size (head + wings) corrections. Even using the small $3 \times 3 \times 3$ \mathbf{k} -mesh, the corrected VBM/CBM energies are very close to the TDL values.

The extrapolated G_0W_0 @PBE band gaps are 1.08 eV and 6.37 eV for Si and cubic BN respectively, in very good agreement with the experimental values (1.17 eV for Si [77], 6.4 eV for BN [78]). For BN, the finite-size-corrected band gap

converges to the TDL much faster than the uncorrected band gap. Using the $7 \times 7 \times 7$ \mathbf{k} -mesh, the G_0W_0 @PBE band gap of cubic BN is 6.57 eV, which is only 0.20 eV larger than the extrapolated TDL value. On the other hand, for Si, the finite-size-corrected band gaps do not converge more quickly with the number of \mathbf{k} -points compared to the uncorrected values. This is mainly due to the surprisingly good error cancellation between the uncorrected VBM and CBM energies. Nevertheless, the corrected G_0W_0 band gap of Si is 1.26 eV using the $7 \times 7 \times 7$ \mathbf{k} -mesh, which is only 0.18 eV larger than the extrapolated TDL value. Overall, we demonstrate that when using the computed finite size corrections, one can obtain reasonably well-converged G_0W_0 excitation energies with respect to the \mathbf{k} -point sampling at an affordable cost.

E. Benchmark of Band Gaps

In this section, we present our G_0W_0 -AC benchmark results for band gaps of 15 semiconductors and rare gas solids. In addition to the 13 solids described in Section III B, we further include ZnO and AlN in the wurtzite structure, marked as wZnO and wAlN. For the 13 cubic semiconductors and rare gas solids, we computed the G_0W_0 @PBE band gaps using \mathbf{k} -meshes ranging from $3 \times 3 \times 3$ to $6 \times 6 \times 6$ without any finite size corrections, then performed linear extrapolations of the form $E(N_{\mathbf{k}}^{-1/3}) = E_{\infty} + aN_{\mathbf{k}}^{-1/3}$ as described in Section III C to obtain the band gaps in the TDL. For wZnO and wAlN, we used $3 \times 3 \times 2$, $4 \times 4 \times 3$ and $6 \times 6 \times 4$ \mathbf{k} -meshes, and then performed linear extrapolations for the G_0W_0 band gaps. We also report the G_0W_0 @PBE band gaps obtained by applying finite size corrections for the $6 \times 6 \times 6$ \mathbf{k} -mesh ($6 \times 6 \times 4$ for wAlN and wZnO). cc-pVTZ/cc-pVTZ-RI basis sets were used for all solids, except for Ne and Ar where we used the aug-cc-pVTZ/aug-cc-pVTZ-RI basis. All lattice constants and detailed band gap values for different \mathbf{k} -meshes can be found in the Supporting Information. The G_0W_0 @PBE band gaps are presented in Table III, and compared to DFT-PBE results and experimental values.

As shown in Table III, G_0W_0 @PBE improves the description of band gaps significantly over PBE, a finding which agrees with previous studies. The mean absolute relative error (MARE) of G_0W_0 @PBE using the \mathbf{k} -point extrapolation is only 5.2% compared to the experimental values. The finite-size-corrected G_0W_0 @PBE using the $6 \times 6 \times 6$ \mathbf{k} -mesh achieves almost the same accuracy (MARE = 5.3%), partly due to the cancellation between the remaining finite size errors and the G_0W_0 @PBE errors (e.g., in MgO and ZnO). Special attention should be placed on wZnO, whose previously reported G_0W_0 band gap values differ by more than 2 eV using different G_0W_0 approximations and codes [74]. A recent benchmark comparison of state-of-the-art full-frequency plane-wave G_0W_0 codes reported that the G_0W_0 @LDA band gap of wZnO extrapolated to the basis set limit is 2.76 eV using a shifted $8 \times 8 \times 5$ \mathbf{k} -mesh [74]. The accurate LAPW+HLOs method in Ref. [76] predicted the G_0W_0 @PBE band gap of wZnO to be 3.01 eV using a $6 \times 6 \times 4$ \mathbf{k} -mesh. Our Gaussian-based G_0W_0 @PBE band gap of wZnO is 3.30 eV using the

TABLE III. Band gaps of semiconductors and rare gas solids from DFT-PBE ($6 \times 6 \times 6$ \mathbf{k} -mesh), finite-size-corrected G_0W_0 @PBE ($6 \times 6 \times 6$ \mathbf{k} -mesh), \mathbf{k} -point extrapolated G_0W_0 @PBE and experiments. The cc-pVTZ basis was used unless otherwise specified. MARE stands for mean absolute relative error compared to the experimental value. All band gap values are in eV.

System	PBE	G_0W_0 ($6 \times 6 \times 6$)	G_0W_0 (extrap.)	Expt.
Si	0.61	1.29	1.08	1.17 [77]
C	4.14	5.91	5.52	5.48 [78]
SiC	1.36	2.59	2.44	2.42 [78]
BN	4.47	6.61	6.41	6.4 [78]
BP	1.42	2.37	2.15	2.4 [78]
wAlN ^a	4.19	6.07	5.89	6.2-6.3 [78]
AIP	1.62	2.56	2.41	2.51 [77]
GaN	1.84	3.40	3.13	3.17 [77]
GaP	1.69	2.45	2.33	2.27 [78]
MgO ^b	4.75	7.73	7.43	7.83 [79]
ZnO	0.94	3.18	2.91	3.4 ^d
wZnO ^a	1.06	3.30	3.08	3.4 [78]
ZnS	2.36	3.90	3.63	3.7 [78]
Ne ^c	11.65	20.05	20.01	21.7 [80]
Ar ^c	8.77	13.33	13.24	14.2 [80]
MARE (%)	42	5.3	5.2	

^a A $6 \times 6 \times 4$ \mathbf{k} -mesh was used for the second (PBE) and third (G_0W_0) columns.

^b The most diffuse p function of Mg was removed to avoid linear dependencies.

^c aug-cc-pVTZ basis.

^d The experimental band gap of zinc blende ZnO is not available, so we use the wurtzite ZnO gap as an approximation.

$6 \times 6 \times 4$ \mathbf{k} -mesh, and the extrapolated TDL band gap is 3.08 eV, in good agreement with these accurate G_0W_0 calculations. Again, we emphasize that only 196 Gaussian basis functions per wZnO unit cell were needed to obtain these results. Based on this benchmark, we can conclude that our all-electron Gaussian-based G_0W_0 implementation is both accurate and efficient when performing G_0W_0 calculations for valence excitations in periodic systems.

F. G_0W_0 for Metals

Metallic systems require additional considerations in G_0W_0 because of the vanishing energy gaps. At the mean-field level, finite temperature smearing is often applied to allow a Fermi-Dirac fractional occupation of the orbitals $\{f_{m\mathbf{k}_m}\}$:

$$f_{m\mathbf{k}_m} = \frac{1}{1 + e^{(\epsilon_{m\mathbf{k}_m} - \mu)/\sigma}}, \quad (54)$$

where μ is the chemical potential and σ is the finite temperature smearing parameter. In the G_0W_0 -AC scheme, Eq. 24 becomes

$$\Pi_{PQ}(\mathbf{q}, i\omega') = \frac{1}{N_{\mathbf{k}}} \sum_{i\mathbf{k}_i a\mathbf{k}_a} v_P^{ia} \frac{(f_{i\mathbf{k}_i} - f_{a\mathbf{k}_a})(\epsilon_{i\mathbf{k}_i} - \epsilon_{a\mathbf{k}_a})}{\omega'^2 + (\epsilon_{i\mathbf{k}_i} - \epsilon_{a\mathbf{k}_a})^2} v_Q^{ai}, \quad (55)$$

TABLE IV. Band energies and bandwidths of bulk Cu calculated by PBE and all-electron Gaussian-based G_0W_0 @PBE using the def2-TZVP basis and $8 \times 8 \times 8$ \mathbf{k} -mesh (with no finite size corrections). Special symmetry points are noted in Fig. 4. The results are compared to G_0W_0 @PBE results using the PAW method in Ref. [21] and G_0W_0 @LDA results using the pseudopotential plane-wave (PPW) scheme in Ref. [81], as well as the experimental values [82].

		PBE	G_0W_0 @PBE	PAW [21]	PPW [81]	Expt. [82]
Positions of d bands	Γ_{12}	-2.18	-2.38	-2.11	-2.81	-2.78
	X_5	-1.46	-1.60	-1.45	-2.04	-2.01
	L_3	-1.60	-1.76	-1.58	-2.24	-2.25
Widths of d bands	$\Gamma_{12} - \Gamma_{25'}$	0.83	0.71	0.69	0.60	0.81
	$X_5 - X_3$	2.97	2.76	2.60	2.49	2.79
	$X_5 - X_1$	3.41	3.18	3.10	2.90	3.17
	$L_3 - L_3$	1.44	1.39	1.26	1.26	1.37
	$L_3 - L_1$	3.42	3.13	3.16	2.83	2.91
Positions of s/p bands	Γ_1	-9.14	-9.05	-9.18	-9.24	-8.60
	$L_{2'}$	-0.80	-0.82	-1.02	-0.57	-0.85
L gap	$L_1^c - L_{2'}$	4.89	5.21	4.98	4.76	4.95

and the sum over states $i\mathbf{k}_i$ and $a\mathbf{k}_a$ runs over all molecular orbitals as all orbitals are partially occupied. This means an electron may be excited within the same energy band, which is termed an intraband transition. ϵ_F in Eq. 27 is set to μ . For metallic systems, intraband transitions lead to a non-vanishing Drude term in the long-wavelength limit, which in principle may be included as a type of finite size correction [21]. However, we have not implemented this correction term here and leave it to future work.

We applied our all-electron Gaussian-based G_0W_0 method to bulk Cu in the fcc structure (lattice constant 3.603 \AA [83]). The band structure of Cu was computed using PBE and G_0W_0 @PBE with the def2-TZVP/def2-TZVP-RIFIT basis [63, 64] and an $8 \times 8 \times 8$ \mathbf{k} -mesh. No finite size corrections were applied here since we have not yet implemented the intraband transition Drude term. A Fermi-Dirac finite temperature smearing ($\sigma = 0.002$ a.u.) was used in the PBE calculation. We note that for metallic systems, one may need more modified Gauss-Legendre grids for stable numerical integration of Eq. 27. For Cu, we used 200 grid points. The results are shown in Fig. 4 and Table IV and compared to previous G_0W_0 studies and experimental values.

As shown in Table IV, PBE predicts larger d bandwidths than the experimental values. For example, the $X_5 - X_1$ and $L_3 - L_1$ bandwidths are 0.24 eV and 0.51 eV too wide. Our Gaussian-based G_0W_0 @PBE approach narrows the d bandwidths by 0.1-0.3 eV, leading to better agreement with the experimental values than PBE. Our G_0W_0 @PBE d bandwidths are also closer to the experimental values (except for $L_3 - L_1$) compared to plane-wave based G_0W_0 results in PAW [21] and pseudopotential PW (PPW) [81] schemes. The PBE positions of the d bands and Γ_1 are predicted to be 0.55-0.65 eV too shallow and 0.54 eV too deep compared to experiment. Our G_0W_0 @PBE corrects these positions by 0.1-0.2 eV, but this is still far from quantitative agreement with experiment. This is similar to the results obtained in the PAW scheme. The much better performance of PPW-based G_0W_0 @LDA for the positions of the d bands should be attributed to the fortuitous error cancellation between the applied pseudopotential and the G_0W_0 approximation, as discussed in Ref. [21]. Finally, our

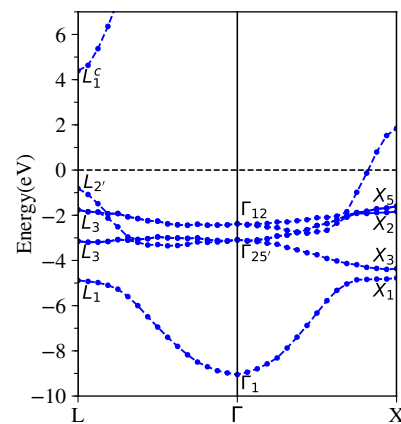


FIG. 4. Band structure of bulk Cu computed by Gaussian-based G_0W_0 @PBE using the def2-TZVP basis and an $8 \times 8 \times 8$ \mathbf{k} -mesh.

G_0W_0 @PBE L gap is 0.26 eV larger than the experimental value, and worse than that obtained from PBE and other G_0W_0 schemes. We believe this is largely due to the missing finite size corrections in our calculation.

IV. CORE EXCITATION BINDING ENERGIES

Core-level X-ray photoelectron spectroscopy (XPS) is a powerful tool for chemical analysis in solids and surfaces [93], because core excitations are sensitive to the atomic environment. A key challenge when using XPS in complex materials is that the assignment of the experimental XPS peaks to the specific atomic sites is often difficult. Thus, it is of great interest to develop accurate first-principles approaches to simulate core excitation binding energies (CEBEs).

One state-of-the-art method for calculating CEBEs is the Δ -self-consistent-field (Δ SCF) approach based on KS-DFT [94], which has been shown to predict accurate CEBEs with approximately 0.2-0.3 eV errors for small molecules [95, 96]. Recently, a restricted open-shell Kohn-Sham (ROKS) ap-

TABLE V. Core excitation binding energies (defined in Eq. 56) calculated by PBE, PBE45 (HF exchange = 45%), all-electron Gaussian-based G_0W_0 @PBE and G_0W_0 @PBE45 (with finite size corrections). The cc-pCVTZ basis set was used for all elements, except for Zn and Ga, where the cc-pVTZ basis was used. $4 \times 4 \times 4$, $4 \times 4 \times 3$ and $4 \times 4 \times 2$ \mathbf{k} -meshes were used for cubic, wurtzite and hexagonal materials respectively. MAE stands for mean absolute error. All energies are in eV.

Core Orbital	Material	PBE	G_0W_0 @PBE	PBE45	G_0W_0 @PBE45	Expt.
C 1s	Diamond	265.66	276.18	280.41	284.48	283.7, 283.9 [78]
	SiC (cubic)	263.36	272.78	277.98	281.57	281.45 [84]
N 1s	BN (hexagonal)	374.20	386.92	391.45	395.37	396.1 [85]
	AlN (wurtzite)	373.44	383.30	390.64	394.30	393.87 [86]
	GaN (wurtzite)	374.75	383.77	392.05	395.50	395.2 [87]
O 1s	BeO (wurtzite)	502.92	514.91	522.38	526.77	527.7 [85]
	MgO (rocksalt)	502.76	512.78	522.26	526.26	527.28 [88]
	ZnO (wurtzite)	504.31	514.83	523.52	527.19	527.45 [86]
Be 1s	BeO (wurtzite)	96.66	104.47	104.17	108.81	109.8 [85]
B 1s	BN (hexagonal)	171.69	183.11	183.18	187.68	188.4 [85]
Al 2p	AlP (cubic)	64.00	68.51	71.09	72.57	72.43 [89]
	AlN (wurtzite)	62.18	67.42	68.48	70.55	70.56 [86]
Si 2p	Si	89.84	94.90	98.26	98.43	98.95 [90]
Mg 2p	MgO (rocksalt)	39.13	43.86	43.63	45.71	46.71 [91], 46.79 [92]
MAE		16.77	8.00	3.59	0.56	

proach [97] has also been proposed which achieves similar accuracy in molecules. However, Δ SCF methods are more complicated to apply to periodic systems, due to the need to treat the Coulomb divergence when one electron is explicitly removed from the system. In practice, some ways to proceed include using finite cluster models [96], adding an excess electron into the conduction bands [98] or using an exact Coulomb cutoff method [99].

GW methods provide an alternative treatment of this problem. By computing the core-state quasiparticle energies through GW , one naturally incorporates relaxation and correlation effects into the core excitation energies [100]. Because GW does not explicitly create a core hole, no special treatment of periodic boundary conditions needs to be used for solids unlike in Δ SCF methods. Recently, Golze *et al.* [37, 101] demonstrated that GW in a Gaussian basis predicts accurate absolute core-level binding energies in XPS with approximately 0.3 eV errors in molecules, when using an eigenvalue self-consistent GW approach or G_0W_0 combined with a density functional with a large percentage of exact exchange. GW in combination with the Bethe-Salpeter equation (GW +BSE) has also been widely used for simulating X-ray absorption (XAS) and X-ray emission (XES) core excitation spectra in molecules and solids [102, 103]. On the other hand, there are relatively few studies that have applied the GW approach to investigate the XPS core excitation energies in solids [24, 25]. Aoki *et al.* [24] showed that G_0W_0 starting from a self-interaction-corrected LDA functional predicted XPS CEBEs within 1 eV of error for a few semiconductors, using a mixed plane-wave and numerical atom-centered orbital basis scheme. Our all-electron Gaussian-based G_0W_0 code is well placed to model core-electron physics, so we apply this method to compute XPS CEBEs in semiconductor materials with explicit periodic boundary conditions.

As discussed in Ref. [37], the frequency structure of the core-state G_0W_0 self-energy is very complicated and cannot be

accurately reproduced by a simple Padé analytic continuation. Thus, in this section, we choose to use the G_0W_0 -CD approach that directly works on the real axis, and always solve the QP equations self-consistently. One more complexity in gapped solids comes from the uncertainty of the Fermi level in experiments. Unlike in molecules, the Fermi level may be anywhere in the gapped region in solids, and may be pinned by possible defects in the system, leading to large variations in the experimental core binding energies [104]. For instance, the reported experimental C 1s core binding energy in diamond varies from 283.25 eV to 291.35 eV [104], making benchmarking of theoretical methods impossible. One way to deal with this issue is to redefine the core excitation binding energy (CEBE) as the difference between the core binding energy and the valence band maximum energy of solids [24]:

$$\text{CEBE} = |e^{\text{core}} - e^{\text{VBM}}|. \quad (56)$$

In this way, the ambiguity in the experimental Fermi level is removed. We thus use this definition of CEBEs to benchmark our G_0W_0 implementation.

We computed 14 CEBEs (the 1s level of C, N, O, Be, B and 2p level of Al, Si, Mg) in typical semiconductors using DFT-PBE and finite-size-corrected G_0W_0 @PBE methods. We also tested the performance of a hybrid functional PBE45 (where 45% of HF exchange is used) and G_0W_0 @PBE45. This choice was suggested in Ref. [101], where the authors found G_0W_0 @PBE45 gave the best G_0W_0 results for core binding energies in molecules. To accurately describe the core states, we used the cc-pCVTZ basis set [105, 106] (which adds extra tight core basis functions to cc-pVTZ) for all elements, except for Zn and Ga where we used the cc-pVTZ basis. The cc-pwCVTZ-RI [64]/cc-pVTZ-RI basis sets were used as auxiliary bases. $4 \times 4 \times 4$, $4 \times 4 \times 3$ and $4 \times 4 \times 2$ \mathbf{k} -meshes were used for cubic, wurtzite and hexagonal materials respectively. Relativistic effects were neglected in this study.

As presented in Table V, CEBEs computed using PBE orbital energies are systematically smaller than the experimental

values by over 16 eV. This large error arises because the KS orbital energies miss the orbital relaxation effects in the final core ionized state. $G_0W_0@PBE$ reduces the errors of PBE significantly, but still has a mean absolute error (MAE) of 8.00 eV. The errors are even larger if only 1s CEBEs are considered. This unsatisfactory performance of $G_0W_0@PBE$ was also observed in molecules [100], where the authors found linearized $G_0W_0@PBE$ gave errors as large as 9 eV in the core binding energies of molecules. Switching to PBE45, the CEBE results are greatly improved over PBE and are even better than $G_0W_0@PBE$, which can be attributed to better error cancellation between the KS core orbital energies and neglecting the orbital relaxation effects after core ionization. Using PBE45 as the starting point, $G_0W_0@PBE45$ further reduces the MAE to only 0.56 eV, and the performance is equally good for the 1s (in C, N, O, Be, B) and 2p (in Al, Si, Mg) CEBEs. We thus conclude that when combined with PBE45, Gaussian-based G_0W_0 shows promise for the accurate simulation of CEBEs in periodic systems.

However, we also notice that the current $G_0W_0@PBE45$ method does not always predict the correct relative CEBEs of the same core orbital in different materials. For example, $G_0W_0@PBE45$ predicts the O 1s CEBE in BeO to be 0.42 eV smaller than in ZnO, while in experiments the O 1s CEBE in BeO is 0.25 eV larger. Such a discrepancy might be partly due to uncertainties in extracting the accurate valence band maximum in XPS experiments [107]. On the computational side, a key candidate for the source of error is the dependence of G_0W_0 on the quality of the starting density functional. As shown in Table V, PBE45 predicts the relative O 1s CEBE shift to be -1.14 eV ($CEBE(BeO) - CEBE(ZnO)$), compared to $+0.25$ eV in experiments. Although $G_0W_0@PBE45$ corrects the PBE45 CEBE shift to -0.42 eV, the result is still of the wrong sign. Therefore, it is interesting to study whether self-consistent GW (e.g., eigenvalue self-consistent GW [101]) can obtain better relative CEBEs for solids in future work.

V. CONCLUSIONS

In this work, we described an all-electron G_0W_0 implementation based on crystalline Gaussian basis sets for periodic systems and benchmarked it on a range of systems including molecules, semiconductors, rare gas solids and metals, for both valence and core excitations. We demonstrated that modern Gaussian bases are an efficient choice for carrying out periodic G_0W_0 calculations, finding that G_0W_0 valence excitation energies are rapidly converged using a small number of basis functions, as seen in the challenging case of ZnO. We developed a finite size correction scheme similar to that used in plane-wave G_0W_0 implementations, allowing our Gaussian-based G_0W_0 calculations to converge to the thermodynamic limit using a moderate amount of \mathbf{k} -point sampling. We also investigated the performance of the G_0W_0 approximation for core excitation binding energies in semiconductors, obtaining promising results in combination with a hybrid density functional with a large fraction of HF exchange. Based on these findings, we conclude that the Gaussian-based G_0W_0 approach is a competitive choice for computing both valence and core excitation energies in weakly-correlated materials. Future work will examine the extension of the current scheme to different types of self-consistency in GW , as well as the combination of G_0W_0 with other quantum chemistry methods, for example using our recent full cell quantum embedding framework [44].

ACKNOWLEDGMENTS

This work was supported by the US Department of Energy, Office of Science under award no. 19390. TZ thanks Zhihao Cui, Yang Gao and Timothy Berkelbach for helpful discussions. Additional support was provided by the Simons Foundation via the Simons Collaboration on the Many Electron Problem, and via the Simons Investigatorship in Physics.

-
- [1] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
 [2] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, *Phys. Rev. Lett.* **49**, 1691 (1982).
 [3] J. P. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983).
 [4] L. J. Sham and M. Schlüter, *Phys. Rev. B* **32**, 3883 (1985).
 [5] J. P. Perdew, W. Yang, K. Burke, Z. Yang, E. K. Gross, M. Scheffler, G. E. Scuseria, T. M. Henderson, I. Y. Zhang, A. Ruzsinszky, H. Peng, J. Sun, E. Trushin, and A. Göring, *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2801 (2017).
 [6] L. Hedin, *Phys. Rev.* **139**, A796 (1965).
 [7] M. S. Hybertsen and S. G. Louie, *Phys. Rev. B* **34**, 5390 (1986).
 [8] F. Aryasetiawan and O. Gunnarsson, *Reports Prog. Phys.* **61**, 237 (1998).
 [9] D. Golze, M. Dvorak, and P. Rinke, *Front. Chem.* **7**, 377 (2019), arXiv:1912.04893.
 [10] X. Ren, P. Rinke, C. Joas, and M. Scheffler, *J. Mater. Sci.* **47**, 7447 (2012), arXiv:1203.5536.
 [11] J. Deslippe, G. Samsonidze, D. A. Strubbe, M. Jain, M. L. Cohen, and S. G. Louie, *Comput. Phys. Commun.* **183**, 1269 (2012), arXiv:1111.4429.
 [12] A. Marini, C. Hogan, M. Grüning, and D. Varsano, *Comput. Phys. Commun.* **180**, 1392 (2009).
 [13] M. Govoni and G. Galli, *J. Chem. Theory Comput.* **11**, 2680 (2015), arXiv:1501.03141.
 [14] L. Martin-Samos and G. Bussi, *Comput. Phys. Commun.* **180**, 1416 (2009).
 [15] M. Schlipf, H. Lambert, N. Zibouche, and F. Giustino, *Comput. Phys. Commun.* **247**, 106856 (2020), arXiv:1812.03717.
 [16] C. Friedrich, S. Blügel, and A. Schindlmayr, *Phys. Rev. B* **81**, 125102 (2010), arXiv:1003.0316.
 [17] H. Jiang, R. I. Gómez-Abal, X. Z. Li, C. Meisenbichler, C. Ambrosch-Draxl, and M. Scheffler, *Comput. Phys. Commun.* **184**, 348 (2013).
 [18] A. Gulans, S. Kontur, C. Meisenbichler, D. Nabok, P. Pavone, S. Rigamonti, S. Sagmeister, U. Werner, and C. Draxl, *J. Phys. Condens. Matter* **26**, 363202 (2014).
 [19] D. Pashov, S. Acharya, W. R. Lambrecht, J. Jackson, K. D. Belashchenko, A. Chantis, F. Jamet, and M. van

- Schilfgaarde, *Comput. Phys. Commun.* **249**, 107065 (2020), [arXiv:1907.06021](#).
- [20] M. Shishkin and G. Kresse, *Phys. Rev. B* **74**, 035101 (2006).
- [21] P. Liu, M. Kaltak, J. Klimeš, and G. Kresse, *Phys. Rev. B* **94**, 165109 (2016), [arXiv:1607.02859](#).
- [22] X. Gonze, B. Amadon, P. M. Anglade, J. M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, T. Deutsch, L. Genovese, P. Ghosez, M. Giantomassi, S. Goedecker, D. R. Hamann, P. Hermet, F. Jollet, G. Jomard, S. Leroux, M. Mancini, S. Mazevet, M. J. Oliveira, G. Onida, Y. Pouillon, T. Rangel, G. M. Rignanese, D. Sangalli, R. Shaltaf, M. Torrent, M. J. Verstraete, G. Zerah, and J. W. Zwanziger, *Comput. Phys. Commun.* **180**, 2582 (2009).
- [23] F. Hüser, T. Olsen, and K. S. Thygesen, *Phys. Rev. B* **87**, 235132 (2013).
- [24] T. Aoki and K. Ohno, *J. Phys. Condens. Matter* **30**, 21LT01 (2018).
- [25] S. Ishii, S. Iwata, and K. Ohno, *Mater. Trans.* **51**, 2150 (2010).
- [26] C. Friedrich, M. C. Müller, and S. Blügel, *Phys. Rev. B* **83**, 081101 (2011).
- [27] M. Stankovski, G. Antonius, D. Waroquiers, A. Miglio, H. Dixit, K. Sankaran, M. Giantomassi, X. Gonze, M. Côté, and G. M. Rignanese, *Phys. Rev. B* **84**, 241201 (2011).
- [28] H. Jiang, *Phys. Rev. B* **97**, 245132 (2018).
- [29] F. Giustino, M. L. Cohen, and S. G. Louie, *Phys. Rev. B* **81**, 115105 (2010), [arXiv:0912.3087](#).
- [30] G. H. Booth, T. Tsatsoulis, G. K.-L. Chan, and A. Grüneis, *J. Chem. Phys.* **145**, 084111 (2016), [arXiv:1603.06457](#).
- [31] X. Blase, C. Attaccalite, and V. Olevano, *Phys. Rev. B* **83**, 115103 (2011), [arXiv:1011.3933](#).
- [32] M. J. Van Setten, F. Weigend, and F. Evers, *J. Chem. Theory Comput.* **9**, 232 (2013).
- [33] F. Bruneval, T. Rangel, S. M. Hamed, M. Shao, C. Yang, and J. B. Neaton, *Comput. Phys. Commun.* **208**, 149 (2016).
- [34] J. Wilhelm, M. Del Ben, and J. Hutter, *J. Chem. Theory Comput.* **12**, 3623 (2016).
- [35] J. Wilhelm, D. Golze, L. Talirz, J. Hutter, and C. A. Pignedoli, *J. Phys. Chem. Lett.* **9**, 306 (2018).
- [36] X. Ren, P. Rinke, V. Blum, J. Wierfink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, *New J. Phys.* **14**, 053020 (2012), [arXiv:1201.0655](#).
- [37] D. Golze, J. Wilhelm, M. J. Van Setten, and P. Rinke, *J. Chem. Theory Comput.* **14**, 4856 (2018).
- [38] P. Koval, M. P. Ljungberg, M. Müller, and D. Sánchez-Portal, *J. Chem. Theory Comput.* **15**, 4564 (2019).
- [39] M. Rohlfing, P. Krüger, and J. Pollmann, *Phys. Rev. B* **48**, 17791 (1993).
- [40] M. Rohlfing, P. Krüger, and J. Pollmann, *Phys. Rev. B* **52**, 1905 (1995).
- [41] J. Wilhelm and J. Hutter, *Phys. Rev. B* **95**, 235123 (2017).
- [42] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. D. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1340 (2018).
- [43] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksson, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Sokolov, and G. K.-L. Chan, (2020), [arXiv:2002.12531](#).
- [44] T. Zhu and G. K.-L. Chan, (2020), [arXiv:2003.01349](#).
- [45] T. Zhu, C. A. Jiménez-Hoyos, J. McClain, T. C. Berkelbach, and G. K.-L. Chan, *Phys. Rev. B* **100**, 115154 (2019).
- [46] Z.-H. Cui, T. Zhu, and G. K.-L. Chan, *J. Chem. Theory Comput.* **16**, 119 (2020), [arXiv:1909.08596](#).
- [47] T. Zhu, Z.-H. Cui, and G. K.-L. Chan, *J. Chem. Theory Comput.* **16**, 141 (2020).
- [48] Q. Sun, T. C. Berkelbach, J. D. McClain, and G. K.-L. Chan, *J. Chem. Phys.* **147**, 164119 (2017), [arXiv:1707.07114](#).
- [49] H. N. Rojas, R. W. Godby, and R. J. Needs, *Phys. Rev. Lett.* **74**, 1827 (1995).
- [50] M. M. Rieger, L. Steinbeck, I. D. White, H. N. Rojas, and R. W. Godby, *Comput. Phys. Commun.* **117**, 211 (1999), [arXiv:9805246 \[cond-mat\]](#).
- [51] R. W. Godby, M. Schlüter, and L. J. Sham, *Phys. Rev. B* **37**, 10159 (1988).
- [52] J. L. Whitten, *J. Chem. Phys.* **58**, 4496 (1973).
- [53] H. J. Vidberg and J. W. Serene, *J. Low Temp. Phys.* **29**, 179 (1977).
- [54] M. J. Van Setten, F. Caruso, S. Sharifzadeh, X. Ren, M. Scheffler, F. Liu, J. Lischner, L. Lin, J. R. Deslippe, S. G. Louie, C. Yang, F. Weigend, J. B. Neaton, F. Evers, and P. Rinke, *J. Chem. Theory Comput.* **11**, 5665 (2015).
- [55] I. Duchemin and X. Blase, *J. Chem. Theory Comput.* **16**, 1742 (2020).
- [56] D. Golze, L. Keller, and P. Rinke, *J. Phys. Chem. Lett.* **11**, 1840 (2020).
- [57] R. Sundararaman and T. A. Arias, *Phys. Rev. B* **87**, 165122 (2013).
- [58] C. Freysoldt, P. Eggert, P. Rinke, A. Schindlmayr, R. W. Godby, and M. Scheffler, *Comput. Phys. Commun.* **176**, 1 (2007), [arXiv:0608215 \[cond-mat\]](#).
- [59] J. Yan, J. J. Mortensen, K. W. Jacobsen, and K. S. Thygesen, *Phys. Rev. B* **83**, 245122 (2011), [arXiv:1104.1273](#).
- [60] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, and T. L. Windus, *J. Chem. Inf. Model.* **59**, 4814 (2019).
- [61] S. Goedecker, M. Teter, and J. Hutter, *Phys. Rev. B* **54**, 1703 (1996).
- [62] P. J. Hay and W. R. Wadt, *J. Chem. Phys.* **82**, 299 (1985).
- [63] F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
- [64] C. Hättig, *Phys. Chem. Chem. Phys.* **7**, 59 (2005).
- [65] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [66] T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- [67] D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **98**, 1358 (1993).
- [68] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, *J. Chem. Phys.* **110**, 7667 (1999).
- [69] B. P. Prascher, D. E. Woon, K. A. Peterson, T. H. Dunning, and A. K. Wilson, *Theor. Chem. Acc.* **128**, 69 (2011).
- [70] N. B. Balabanov and K. A. Peterson, *J. Chem. Phys.* **123**, 064107 (2005).
- [71] F. Weigend, A. Köhn, and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2002).
- [72] J. G. Hill and J. A. Platts, *J. Chem. Phys.* **128**, 044104 (2008).
- [73] R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- [74] T. Rangel, M. Del Ben, D. Varsano, G. Antonius, F. Bruneval, F. H. da Jornada, M. J. van Setten, O. K. Orhan, D. D. O'Regan, A. Canning, A. Ferretti, A. Marini, G. M. Rignanese, J. Deslippe, S. G. Louie, and J. B. Neaton, *Comput. Phys. Commun.* , 107242 (2020).

- [75] B. C. Shih, Y. Xue, P. Zhang, M. L. Cohen, and S. G. Louie, *Phys. Rev. Lett.* **105**, 146401 (2010).
- [76] H. Jiang and P. Blaha, *Phys. Rev. B* **93**, 115203 (2016).
- [77] O. Madelung, *Semiconductors: data handbook*, 3rd ed. (Springer-Verlag New York, 2004).
- [78] T. Chiang, K. Frank, H. Freund, A. Goldmann, F. J. Himpsel, U. Karlsson, R. Lecky, and W. Schneider, *Electronic structure of solids: Photoemission spectra and related data* (Springer, 1989).
- [79] R. C. Whited, C. J. Flaten, and W. C. Walker, *Solid State Commun.* **13**, 1903 (1973).
- [80] N. Schwentner, F. J. Himpsel, V. Saile, M. Skibowski, W. Steinmann, and E. E. Koch, *Phys. Rev. Lett.* **34**, 528 (1975).
- [81] A. Marini, G. Onida, and R. Del Sole, *Phys. Rev. Lett.* **88**, 4 (2002).
- [82] R. Courths and S. Hüfner, *Phys. Rep.* **112**, 53 (1984).
- [83] L. Schimka, J. Harl, and G. Kresse, *J. Chem. Phys.* **134**, 24116 (2011).
- [84] J. R. Waldrop and R. W. Grant, *Appl. Phys. Lett.* **56**, 557 (1990).
- [85] K. Hamrin, G. Johansson, U. Gelius, C. Nordling, and K. Siegbahn, *Phys. Scr.* **1**, 277 (1970).
- [86] T. D. Veal, P. D. King, S. A. Hatfield, L. R. Bailey, C. F. McConville, B. Martel, J. C. Moreno, E. Frayssinet, F. Semond, and J. Zúñiga-Pérez, *Appl. Phys. Lett.* **93**, 202108 (2008).
- [87] T. L. Duan, J. S. Pan, and D. S. Ang, *Appl. Phys. Lett.* **102**, 201604 (2013).
- [88] R. K. Chellappan, Z. Li, and G. Hughes, *Appl. Surf. Sci.* **285**, 153 (2013).
- [89] J. R. Waldrop, R. W. Grant, and E. A. Kraut, *J. Vac. Sci. Technol. B Microelectron. Nanom. Struct.* **11**, 1617 (1993).
- [90] E. T. Yu, E. T. Croke, T. C. McGill, and R. H. Miles, *Appl. Phys. Lett.* **56**, 569 (1990).
- [91] X. Liu, X. Wang, D. Wang, J. Tang, X. Fang, D. Fang, Y. Li, B. Yao, X. Ma, H. Wang, and Z. Wei, *Vacuum* **134**, 136 (2016).
- [92] H. S. Craft, R. Collazo, M. D. Losego, S. Mita, Z. Sitar, and J. P. Maria, *J. Appl. Phys.* **102**, 074104 (2007).
- [93] C. S. Fadley, *J. Electron Spectros. Relat. Phenomena* **178-179**, 2 (2010).
- [94] N. A. Besley, A. T. Gilbert, and P. M. Gill, *J. Chem. Phys.* **130**, 124308 (2009).
- [95] N. Pueyo Bellafont, F. Viñes, and F. Illas, *J. Chem. Theory Comput.* **12**, 324 (2016).
- [96] J. M. Kahk and J. Lischner, *Phys. Rev. Mater.* **3**, 100801 (2019), [arXiv:1904.04823](https://arxiv.org/abs/1904.04823).
- [97] D. Hait and M. Head-Gordon, *J. Phys. Chem. Lett.* **11**, 775 (2020), [arXiv:1912.05249](https://arxiv.org/abs/1912.05249).
- [98] E. Pehlke and M. Scheffler, *Phys. Rev. Lett.* **71**, 2338 (1993).
- [99] T. Ozaki and C. C. Lee, *Phys. Rev. Lett.* **118**, 026401 (2017).
- [100] M. J. Van Setten, R. Costa, F. Viñes, and F. Illas, *J. Chem. Theory Comput.* **14**, 877 (2018).
- [101] D. Golze, L. Keller, and P. Rinke, *J. Phys. Chem. Lett.* **11**, 1840 (2020).
- [102] J. Vinson, J. J. Rehr, J. J. Kas, and E. L. Shirley, *Phys. Rev. B* **83**, 115106 (2011), [arXiv:1010.0025](https://arxiv.org/abs/1010.0025).
- [103] K. Gilmore, J. Vinson, E. L. Shirley, D. Prendergast, C. D. Pemmaraju, J. J. Kas, F. D. Vila, and J. J. Rehr, *Comput. Phys. Commun.* **197**, 109 (2015), [arXiv:1601.07728](https://arxiv.org/abs/1601.07728).
- [104] M. Walter, F. Mangolini, J. B. McClimon, R. W. Carpick, and M. Moseler, (2019), [arXiv:1902.02958](https://arxiv.org/abs/1902.02958).
- [105] D. E. Woon and T. H. Dunning, *J. Chem. Phys.* **103**, 4572 (1995).
- [106] K. A. Peterson and T. H. Dunning, *J. Chem. Phys.* **117**, 10548 (2002).
- [107] S. A. Chambers, T. Droubay, T. C. Kaspar, and M. Gutowski, *J. Vac. Sci. Technol. B* **22**, 2205 (2004).