


## Efficient core-excited state orbital perspective on calculating x-ray absorption transitions in determinant framework

Subhayan Roychoudhury<sup>✉\*</sup> and David Prendergast<sup>†</sup>

*The Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

 (Received 23 August 2022; revised 24 November 2022; accepted 19 January 2023; published 26 January 2023)

X-ray absorption spectroscopy (XAS) is an explicit probe of the unoccupied electronic structure of materials and an invaluable tool for fingerprinting various electronic properties and phenomena. Computational methods capable of simulating and analyzing such spectra are therefore in high demand for complementing experimental results and for extracting valuable insights therefrom. In particular, a recently proposed first-principles approach titled many-body XAS (MBXAS) [Y. Liang *et al.*, *Phys. Rev. Lett.* **118**, 096402 (2017)], which approximates the final (initial) state as a Slater determinant constructed from Kohn-Sham (KS) orbitals optimized in the absence (presence) of the relevant core electron has shown promising prospects in evaluating the transition amplitudes. In this article, we show that the MBXAS approach can be rederived using a transition operator expressed entirely in the basis of core-excited state KS orbitals and that this reformulation offers substantial practical and conceptual advantages. In addition to circumventing previous issues of convergence with respect to the number of unoccupied ground-state orbitals, the aforementioned representation reduces the computational expense by rendering the calculation of such orbitals unnecessary altogether. The reformulated approach also provides a direct pathway for comparing the many-body approximation with the so-called single-particle treatment and indicates the relative importance in observed XAS intensity of the relaxation of the valence occupied subspace induced by the core excitation. Finally, using the core-excited state basis, we define auxiliary orbitals for x-ray absorption and demonstrate their utility in explaining the spectral intensity by contrasting them with single-particle approximations to the excited state.

DOI: [10.1103/PhysRevB.107.035146](https://doi.org/10.1103/PhysRevB.107.035146)

### I. INTRODUCTION

Several major aspects of modern technology are heavily reliant on our ability to investigate and engineer the electronic structure of functional materials. Therefore, a considerable portion of the advancements in electronics and computing [1–3], energy storage [4,5], renewable energy harvesting [6,7], drug discovery [8–10], discovery of magnetic materials [11], etc., can be attributed to the advancements in the theoretical and experimental research on electronic structure theory. On the experimental front, the electronic structure of materials is usually explored by creating electronic excitations. Among such methods, x-ray absorption spectroscopy (XAS) [12–19], which offers an element-specific probe, has emerged as one of the most popular and effective techniques over the last few decades. Besides providing a relatively straightforward avenue for investigating, explicitly, the electronic structure (mostly of the unoccupied subspace) of materials, the XAS process is widely employed for fingerprinting various chemical attributes and phenomena of interest [20–24]. Within a quasielectron framework, the XAS process amounts to excitation of the core electrons of the desired atomic species to the unoccupied levels and measures, albeit typically through indirect means, the cross section of such transitions.

As with most experimental techniques, the interpretation of XAS data and the extraction of information related to electronic structure and properties therefrom is often facilitated by complementary theoretical and computational research [25]. Development of efficient and accurate theoretical models for simulating XAS is, therefore, an active and dynamic field of research.

Among the various first-principles methods employed in computational electronic structure theory, Kohn-Sham (KS) density-functional theory (DFT) has witnessed unprecedented popularity in a myriad of applications owing to its favorable system-size scaling, modest computational requirements, and typically reasonable accuracy. Consequently, in the field of x-ray spectroscopy, substantial effort has been invested in developing and applying theoretical models built upon KS DFT. For instance, the response-based approaches, such as linear-response (LR) time-dependent DFT (TDDFT) [26–29] or the Bethe-Salpeter equation (BSE) [30–32], aim to evaluate the absorption cross section from the response of the electron density to a change in the potential. Such response, in turn, is obtained most commonly by approximating the excited states as a linear combination of singles excitations of the KS ground state (GS). However, the response-based technique can prove to be expensive, especially when it involves explicit solution of the relevant equation (LR-TDDFT/BSE) within an all-electron calculation [31,32]. Additionally, the response-based approach, within the commonly used adiabatic approximation, can incur substantial errors in core-level spectroscopy,

\*roychos@tcd.ie

†dgpendergast@lbl.gov

due to the inadequate description of the electron-hole (e-h) interaction [28,33]. A viable inexpensive alternative is the constrained-occupation approach, whereby the excited state is modeled by maintaining a hole in the core system with the help of additional constraints. Interestingly, such a treatment can be used in a pseudopotential-based calculation by representing the core of the excited atom by a modified pseudopotential that encodes the effects of the core hole [14,34,35]. The approximate KS orbitals corresponding to the core-excited state can then be obtained from a self-consistent field (SCF) calculation employing the aforementioned core-hole pseudopotential either within the full core-hole (FCH) treatment, which minimizes the energy of the core-ionized system with a net positive charge, or within the excited-electron and core-hole (XCH) treatment, which minimizes the energy of the neutral core-excited state.

Within the so-called single-particle approximation, one can estimate the absorption amplitude as that of a transition of a noninteracting electron from the core orbital to the relevant unoccupied KS orbital. Such a simplified calculation, however, neglects the explicit effects of the relaxation of the occupied orbitals in response to the core excitation, resulting often in severe inaccuracies in the absorption intensities. To mitigate these errors, the recently proposed many-body x-ray absorption spectroscopy (MBXAS) method [35,36] relies on two separate SCF calculations:

(1) a GS calculation generating the orbitals for the initial state, and

(2) a FCH/XCH calculation generating the orbitals for the core-excited state.

Each state is then expressed as a Slater determinant (SD) composed of the corresponding orbitals and the transition amplitude is obtained, typically within the dipole approximation, as a many-body matrix element of the transition operator between the initial and the final state SDs. This determinant-based technique, which only requires simple pseudopotential-based KS DFT calculations and consequently retains their low computational cost and scaling, has been shown to significantly improve the peak intensities of the single-particle approximation, especially for the O *K* edge of transition metal oxides, and to obtain very good agreement with the experimental spectra.

Building upon this development, in this article we reformulate the MBXAS formalism by recasting the relevant matrix elements with respect to the core-hole basis (CHB) which improves the computational efficiency, eliminates issues of convergence, and enables direct comparison with the single-particle approximation. Additionally, we develop the concept of auxiliary orbitals for XAS which helps with the association of physically motivated single-particle orbitals to the excited many-electron system.

The rest of the article is organized as follows. We introduce the reformulated treatment of MBXAS in the next section, which is followed by a prescription, inspired from Ref. [36], for achieving computational simplification. The following subsection, where we present and discuss the results of our calculation on some representative systems, starts with a discussion, along with some illustrative examples, on the computational advantages of recasting the expressions to the core-hole basis set. This is then followed by a separation of

the terms contributing to the MBXAS amplitude into categories of different physical origins which facilitate a direct comparison with the single-particle amplitude, as discussed in the next subsection. Finally, we move on to the discussion of auxiliary orbitals corresponding to x-ray excitation.

## II. METHOD

### A. General background

Within a quasiparticle-inspired framework, x-ray absorption results in the neutral excitation of an electronic system typically from its ground state to an excited final state containing a core hole and a conduction electron. In order to simplify the terminology, without loss of generality, in the following, the term ‘‘core’’ is reserved only to denote the electron/hole corresponding to the specific orbital emptied as a consequence of x-ray absorption. Denoting by  $a_i^\dagger$  ( $a_0^\dagger$ ) the creation operator corresponding to the  $i$ th valence level  $|\phi_i\rangle$  (core level  $|\phi_0\rangle$ ), the ground state (GS) of the system can be expressed as the Slater determinant (SD)

$$|\Psi_{\text{GS}}\rangle = \left( \prod_{i=1}^N a_i^\dagger \right) a_0^\dagger |\emptyset\rangle, \quad (1)$$

where  $|\emptyset\rangle$  denotes the null state. In contrast, a final state can be constructed by populating the KS orbitals obtained from a SCF calculation performed in absence of the relevant core electron, which, in accordance with the standard convention, will be called a full core-hole (FCH) SCF hereafter. Using a  $\sim$  to indicate quantities corresponding to such a core-excited SCF, the core-ionized system in its lowest energy state will be given by

$$|\Psi_{\text{FCH}}\rangle = \left( \prod_{\mu=1}^N \tilde{a}_\mu^\dagger \right) |\emptyset\rangle, \quad (2)$$

with the explicit absence of the core orbital, indexed as 0. Then, within the singles approximation, any neutral final state  $|\Psi_f\rangle$  in which, in addition to the lowest  $N$  valence orbitals, the conduction orbital  $|\tilde{\phi}_f\rangle = \tilde{a}_f^\dagger |\emptyset\rangle$  is now occupied, can be constructed from  $|\Psi_{\text{FCH}}\rangle$  as

$$|\Psi_f\rangle = |\Psi_{\text{FCH}}^{+f}\rangle = \tilde{a}_f^\dagger |\Psi_{\text{FCH}}\rangle. \quad (3)$$

The MBXAS approximation [35,36] then entails computing the frequency ( $\omega$ ) dependent absorption cross section as

$$\sigma(\omega) \propto \omega \sum_f |\langle \Psi_f | \hat{O} | \Psi_{\text{GS}} \rangle|^2 \delta(E_f - E_{\text{GS}} - \hbar\omega), \quad (4)$$

where  $\hat{O}$  is the many-body transition operator (used within the dipole approximation here) and  $E_{\text{GS}}$  and  $E_f$  are the total energies of the ground and the final state, respectively. The operator  $\hat{O}$  essentially governs the electronic transitions and, for an  $N$ -electron system, can be expressed as a sum of single-particle operators,

$$\hat{O} = \sum_{i=N}^N \hat{o}_i, \quad (5)$$

where, within the dipole approximation, the transition-operator for the  $i$ th electron can be expressed as  $\hat{o}_i = \epsilon \cdot \hat{\mathbf{r}}_i$ ,

where  $\epsilon$  denotes the polarization of the incident photon and  $\hat{\mathbf{r}}_i$  is the position operator for the  $i$ th electron.

$E_{\text{GS}}$  is obtained from the GS SCF calculation.  $E_f$ , on the other hand, can be approximated from an FCH calculation as

$$E_f = E_{\text{FCH}} + \tilde{\epsilon}_f, \quad (6)$$

where  $\tilde{\epsilon}_f$  is the KS eigenvalue of the conduction orbital  $|\tilde{\phi}_f\rangle$ . Additionally, a constant rigid shift, obtained purely from considerations of total energies of the system and of the core-excited atom (in their respective ground and the lowest core excited states) [21,37,38], can be applied across all  $E_f$  to align the spectra.

We note here that our single-determinant ansatz for the many-body wave functions and the associated single-configuration prescription for how we calculate the individual electronic orbital eigenvalues, through constrained orbital occupation, naturally limit the MBXAS approach to systems with simpler (ground and excited) states. Specifically, we would expect MBXAS calculations to describe  $K$ -edge (i.e.,  $1s$ ) excitations well, particularly for lighter elements (i.e., soft x rays), where both ground and core-excited states may be well approximated by a single configuration, indicating which orbitals are occupied and which are not. However, for edges accessing core orbitals with higher angular momentum, e.g.,  $L_{2,3}$  edges ( $2p$  excitations), the core-excited state involves interactions between multiple nearly degenerate core and valence (e.g.,  $3d$ ) orbital manifolds, giving rise to so-called multiplet effects [39] which necessitate a multiconfigurational description of the excited states and would render the MBXAS approach, as written, inaccurate. In addition, the core-excited state energies, in systems with anisotropic dielectric response functions, may depend on the particular excited core orbital orientation and require separate SCF calculations for each orientation-dependent excited state. This is not a concern for (spherical)  $s$  excitations typically explored using MBXAS.

### B. MBXAS in core-excited state basis

In the original formulation [35,36], the transition operator was represented in terms of the GS orbitals as

$$\hat{O} = \sum_{i,j} \langle \phi_i | \hat{o} | \phi_j \rangle a_i^\dagger a_j, \quad (7)$$

with  $\hat{o}$  standing for the single-particle transition operator and the sum extending over all pairs of orbitals of the GS system. The MBXAS expressions obtained using such representation [the relevant transition amplitudes are presented later in Eqs. (23) and (24)] will, hereafter, be referred to as the ground-state basis (GSB) MBXAS in order to contrast it with the CHB-MBXAS treatment which is central to this work.

As with any operator,  $\hat{O}$  can, in fact, be expanded in any complete single-particle Hilbert space representation. This prompts us to express the transition operator with respect to the FCH state orbitals as

$$\begin{aligned} \hat{O} &= \sum_{l,h} \langle \tilde{\phi}_l | \hat{o} | \tilde{\phi}_h \rangle \tilde{a}_l^\dagger \tilde{a}_h \\ &= \sum_{l,h} \tilde{o}_{lh} \tilde{a}_l^\dagger \tilde{a}_h, \end{aligned} \quad (8)$$

which, as discussed in the following sections, offers substantial advantage in terms of computation and interpretation of the x-ray absorption spectra. By combining Eqs. (1), (2), (3), and (8), we obtain the expression for the transition amplitude

$$\begin{aligned} A_{\text{GS} \rightarrow f} &= \langle \Psi_f | \hat{O} | \Psi_{\text{GS}} \rangle \\ &= \sum_{l,h} \tilde{o}_{lh} \langle \emptyset | \left( \prod_{\mu=1}^N \tilde{a}_\mu \right) \tilde{a}_f \tilde{a}_l^\dagger \tilde{a}_h \left( \prod_{i=1}^N a_i^\dagger \right) a_0^\dagger | \emptyset \rangle \\ &= \sum_{l=1 \dots N, f} \tilde{o}_l \langle \emptyset | \left( \prod_{\mu=1}^N \tilde{a}_\mu \right) \tilde{a}_f \tilde{a}_l^\dagger \tilde{a}_0 \left( \prod_{i=1}^N a_i^\dagger \right) a_0^\dagger | \emptyset \rangle \\ &= \sum_{l=1 \dots N, f} \tilde{o}_l \langle \Psi_f | \tilde{a}_l^\dagger \tilde{a}_0 | \Psi_{\text{GS}} \rangle, \end{aligned} \quad (9)$$

where, for the sake of simplicity, we have used the notation  $\tilde{o}_l = \tilde{o}_{l0}$ . The third line follows from the observation that, in the sum over  $h$ , all terms except for that corresponding to  $h = 0$  (i.e., where  $h$  is essentially the core level) vanish. The approximation  $|\phi_0\rangle \approx |\tilde{\phi}_0\rangle$  is justified due to the extremely localized nature of the core orbital [33].

Now, for any  $f > N$  and  $l$  in the range  $1 \dots N$ , we introduce the many-electron state

$$|\Psi_{\text{FCH}}^{+f-l+0}\rangle = (-1)^{\gamma_l} (\tilde{a}_0^\dagger \tilde{a}_l | \Psi_f \rangle), \quad (10)$$

which has an index-ordered set of occupied orbitals, where the factor  $(-1)^{\gamma_l}$ , with  $\gamma_l = (2N + 1 - l)$ , originates from the anticommutation relations of the creation and annihilation operators. Note that the inclusion of this factor is necessary to account for index reordering of the operators. More specifically, in any expression involving the creation and annihilation operators, an exchange of two consecutive operators generates a factor of  $-1$ . Thus, given that, in order to obtain  $|\Psi_{\text{FCH}}^{+f-l+0}\rangle$  from  $|\Psi_f\rangle$ , the operators  $\tilde{a}_l$  and  $\tilde{a}_0^\dagger$  must be moved  $N + 1 - l$  and  $N$  places, respectively, a resultant factor of  $(-1)^{2N+1-l}$  is introduced (more details on this can be found in Appendix A). Equation (9) can then be rewritten as

$$\begin{aligned} A_{\text{GS} \rightarrow f} &= \sum_{l=1 \dots N} (-1)^{\gamma_l} \tilde{o}_l \langle \Psi_{\text{FCH}}^{+f-l+0} | \Psi_{\text{GS}} \rangle \\ &\quad + (-1)^{\gamma_f} \tilde{o}_f \langle \Psi_{\text{FCH}}^{+0} | \Psi_{\text{GS}} \rangle, \end{aligned} \quad (11)$$

where  $\gamma_f = N$ , and we have used the notation  $|\Psi_{\text{FCH}}^{+f-f+0}\rangle = |\Psi_{\text{FCH}}^{+0}\rangle = (-1)^N \tilde{a}_0^\dagger \tilde{a}_f | \Psi_f \rangle$ . Now, noting that an overlap of two SDs can be expressed as a determinant composed of the constituent orbital overlaps, the many-body overlap in the second term in Eq. (11) can be expressed as the determinant of an  $N \times N$  matrix of valence orbital overlaps,

$$\langle \Psi_{\text{FCH}}^{+0} | \Psi_{\text{GS}} \rangle = \det \begin{bmatrix} \xi_{1,1}^* & \cdots & \xi_{1,N}^* \\ \vdots & \ddots & \vdots \\ \xi_{N,1}^* & \cdots & \xi_{N,N}^* \end{bmatrix}, \quad (12)$$

where  $\xi_{i,j} = \langle \phi_j | \tilde{\phi}_i \rangle$  denotes the overlap between the GS and the FCH KS orbitals and the core-orbital overlap is omitted since we approximate that it remains unchanged in both SCFs (this approximation is seen to be valid to a high degree of accuracy [33]).

The first term in Eq. (11) contains the determinant of a modified  $N \times N$  valence overlap matrix,  $\langle \Psi_{\text{FCH}}^{+f-l+0} | \Psi_{\text{GS}} \rangle$ , where the  $l$ th row has been removed and a new row  $f$  added:

$$\langle \Psi_{\text{FCH}}^{+f-l+0} | \Psi_{\text{GS}} \rangle = \det \begin{bmatrix} \xi_{1,1}^* & \cdots & \xi_{1,N}^* \\ \vdots & \ddots & \vdots \\ \xi_{l-1,1}^* & \cdots & \xi_{l-1,N}^* \\ \xi_{l+1,1}^* & \cdots & \xi_{l+1,N}^* \\ \vdots & \ddots & \vdots \\ \xi_{N,1}^* & \cdots & \xi_{N,N}^* \\ \xi_{f,1}^* & \cdots & \xi_{f,N}^* \end{bmatrix}. \quad (13)$$

Thus, with the help of Eqs. (11), (12), and (13), the MBXAS method approximates the transition amplitude for any final state  $|\Psi_{\text{FCH}}^{+f}\rangle$  from the results of two converged SCF calculations of KS DFT: one for the ground state and one for the core-ionized state.

### C. Computational simplification

Now, we introduce the notation

$$\alpha_q = [\xi_{q,1} \xi_{q,2} \cdots \xi_{q,N}], \quad (14)$$

for an  $N$ -dimensional vector of valence orbital overlaps, and the determinant in Eq. (12) can be rewritten as

$$A_{\text{det}} = \langle \Psi_{\text{GS}} | \Psi_{\text{FCH}}^{+0} \rangle = \det \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}, \quad (15)$$

which can be reexpressed using exterior product algebra as

$$A_{\text{det}} = \alpha_1 \wedge \cdots \wedge \alpha_N, \quad (16)$$

where the ordered exterior product  $\alpha_1 \wedge \cdots \wedge \alpha_N$  is, in fact, one-dimensional [40].

Assuming that this determinant is nonzero, then the  $N$  vectors  $\alpha_q$  define a linearly independent set that spans this  $N$ -dimensional space. Now, using the exterior-product identity, the complex-conjugate of the determinantal overlap of Eq. (13) can then be rewritten as

$$\langle \Psi_{\text{GS}} | \Psi_{\text{FCH}}^{+f-l+0} \rangle = \alpha_1 \wedge \cdots \wedge \alpha_{l-1} \wedge \alpha_{l+1} \wedge \cdots \wedge \alpha_N \wedge \alpha_f,$$

and, since  $f > N$ ,  $\alpha_f$  must be linearly dependent on the  $N$  vectors,  $\alpha_1, \dots, \alpha_N$ . This linear dependence leads to

$$\alpha_f = \sum_{k=1}^N \kappa_{f,k} \alpha_k, \quad (17)$$

with  $\{\kappa_{f,k}\}$  denoting a set of unknown coefficients which can be computed by expressing the linear relations of Eq. (17), for  $f$  in the range  $N+1, \dots, M$ , in the form of a matrix equation as

$$\begin{bmatrix} \alpha_{N+1} \\ \vdots \\ \alpha_M \end{bmatrix} = \begin{bmatrix} \kappa_{N+1,1} & \cdots & \kappa_{N+1,N} \\ \vdots & \ddots & \vdots \\ \kappa_{M,1} & \cdots & \kappa_{M,N} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix}. \quad (18)$$

Denoting the first, second, and third matrices in Eq. (18) as  $A'_{\text{mat}}$ ,  $K_{\text{mat}}$ , and  $A_{\text{mat}}$ , respectively,  $K_{\text{mat}}$  can be calculated from

the equation

$$K_{\text{mat}} = A'_{\text{mat}} \cdot (A_{\text{mat}})^{-1}. \quad (19)$$

Now, the determinantal overlap of Eq. (13) can be further rewritten as

$$\begin{aligned} \langle \Psi_{\text{GS}} | \Psi_{\text{FCH}}^{+f-l+0} \rangle &= \alpha_1 \wedge \cdots \wedge \alpha_{l-1} \wedge \alpha_{l+1} \wedge \cdots \wedge \alpha_N \wedge \left( \sum_{k=1}^N \kappa_{f,k} \alpha_k \right) \\ &= (-1)^{N-l} \kappa_{f,l} (\alpha_1 \wedge \cdots \wedge \alpha_N) \\ &= (-1)^{N-l} \kappa_{f,l} A_{\text{det}}, \end{aligned} \quad (20)$$

where we have used Eq. (17) and the anticommutative properties of the exterior product:

$$\alpha_i \wedge \alpha_j = -\alpha_j \wedge \alpha_i \text{ and } \alpha_i \wedge \alpha_i = 0. \quad (21)$$

Then, from Eq. (11), the transition amplitude can be calculated as

$$\langle \Psi_f | \hat{O} | \Psi_{\text{GS}} \rangle = (-1)^N \left[ \tilde{\delta}_f - \sum_{l=1..N} \kappa_{f,l}^* \tilde{\delta}_l \right] A_{\text{det}}^*. \quad (22)$$

The parameters used in our DFT-based calculations are discussed in Appendix B.

## III. DISCUSSIONS

### A. Computational advantage

It is instructive to compare the MBXAS formalism in the CHB approach as presented above with the original GSB-MBXAS proposed in Ref. [36]. Both of the approaches aim to compute the transition amplitude from the identical expression  $A_{\text{GS} \rightarrow f} = \langle \Psi_f | \hat{O} | \Psi_{\text{GS}} \rangle$  with the states  $|\Psi_{\text{GS}}\rangle$  and  $|\Psi_f\rangle$  being given by Eq. (1) and Eq. (3), respectively. Therefore, at convergence, the spectra calculated using the CHB and the GSB approaches should, in principle, be the same. In practice, however, there can be substantial differences.

In the GSB treatment, the transition amplitude within the singles approximation is simplified to

$$\begin{aligned} A_{\text{GS} \rightarrow f} &= \sum_{c>N} \langle \Psi_{\text{FCH}}^{+f} | \Psi_{\text{GS}}^{+c-0} \rangle \langle \Psi_{\text{GS}}^{+c-0} | \hat{O} | \Psi_{\text{GS}} \rangle \\ &= \sum_{c>N} o_c \langle \Psi_{\text{FCH}}^{+f} | \Psi_{\text{GS}}^{+c-0} \rangle, \end{aligned} \quad (23)$$

where  $o_c = \langle \phi_c | \hat{O} | \phi_0 \rangle$  and

$$|\Psi_{\text{GS}}^{+c-0}\rangle = (-1)^N a_c^\dagger a_0 | \Psi_{\text{GS}} \rangle. \quad (24)$$

The relevant many-body overlap is then given by

$$\langle \Psi_{\text{FCH}}^{+f} | \Psi_{\text{GS}}^{+c-0} \rangle = (-1)^N \begin{bmatrix} \xi_{1,1}^* & \cdots & \xi_{1,N}^* & \xi_{1,c}^* \\ \vdots & \ddots & \vdots & \vdots \\ \xi_{N,1}^* & \cdots & \xi_{N,N}^* & \xi_{N,c}^* \\ \xi_{f,1}^* & \cdots & \xi_{f,N}^* & \xi_{f,c}^* \end{bmatrix}. \quad (25)$$

Note that, in principle, the sum in Eq. (23) extends over all unoccupied GS orbitals—bound and unbound both. Thus, if one fails to include any GS unoccupied orbital which has a nonzero overlap with any of the valence occupied FCH-state

orbitals (i.e.,  $|\tilde{\phi}_1\rangle, \dots, |\tilde{\phi}_N\rangle$ ), then, in principle,  $A_{GS \rightarrow f}$  will be under-converged for any  $f > N$ . Additionally, if, for a given  $f > N$ ,  $|\tilde{\phi}_f\rangle$  overlaps with any missing unoccupied GS orbitals, then the corresponding  $A_{GS \rightarrow f}$  will be under-converged. Therefore, convergence in Eq. (23) can be particularly difficult to achieve if one or more of these FCH-state orbitals has appreciable overlap with the unbound unoccupied orbitals of the GS [41].

As an aside, we may note that similar convergence issues should plague any computational method for simulating XAS that is rooted in a GS orbital representation of the XAS transition amplitude. Prominent examples include both the aforementioned LR-TDDFT [26–29] and BSE [30–32] approaches, which are both expressed within a GS orbital basis that includes unoccupied orbitals.

As illustrative examples, in Fig. 1, we present comparative studies of the GSB and the CHB treatments for the C  $K$ -edge spectrum of methanol ( $\text{CH}_3\text{OH}$ ), O  $K$ -edge spectrum of acetone ( $\text{C}_3\text{H}_6\text{O}$ ), and the O  $K$ -edge spectrum of cupric oxide ( $\text{CuO}$ ). The GSB spectra are seen to be heavily dependent on the number of unoccupied GS orbitals employed in the calculation. The fact that the large GSB spectra do not coincide with the scaled small GSB counterparts indicates that the under-converged GSB spectra differ inherently in their line shapes.

Since the sum in Eq. (9) is limited to a finite number ( $N + 1$ ) of terms anyway, this complication is circumvented entirely in the CHB treatment as evident from Fig. 1, which shows, as expected, that the CHB spectra do not depend on the number of orbitals included in the calculation, as long as all FCH-state orbitals within the desired energy range are included. In addition to not suffering from convergence issues, for any given energy range of the spectrum, the CHB treatment is computationally favorable over the GSB treatment since, unlike the latter, the former does not require the computation of any unoccupied GS orbital.

For a simulation involving  $N_{\text{tot}}$  number of GS orbitals (including the valence occupied orbitals as well as the unoccupied ones), within the GSB approach, a practical MBXAS calculation requires the execution of the following steps separately for each of the GS and the FCH systems:

- (1) Calculate the SCF electron density.
- (2) Calculate the unoccupied KS orbitals non-self-consistently (NSCF).
- (3) Transform the KS orbitals into an optimal basis representation to facilitate faster Brillouin zone integration.

In principle, step 3 is optional for the MBXAS procedure. Additionally, for the GS system, we

- (4) calculate the single-particle transition matrix elements  $o_c$  for all unoccupied GS orbitals  $|\phi_c\rangle$ .

Following this, we

- (5) calculate the ( $N_{\text{tot}} \times N_{\text{tot}}$ ) overlap matrix between the ground and the FCH state orbitals.

Finally, we

- (6) compute the MBXAS amplitudes from Eq. (23).

On the other hand, in the CHB treatment, for the GS system, step 2 is not required at all while step 3 can become substantially cheaper due to the absence of the unoccupied orbitals. The same is true for step 5 which now calculates the ( $N \times N_{\text{tot}}$ ) overlap matrix  $\xi_{i,j} \forall i \leq N_{\text{tot}}$  and  $j \leq N$  and

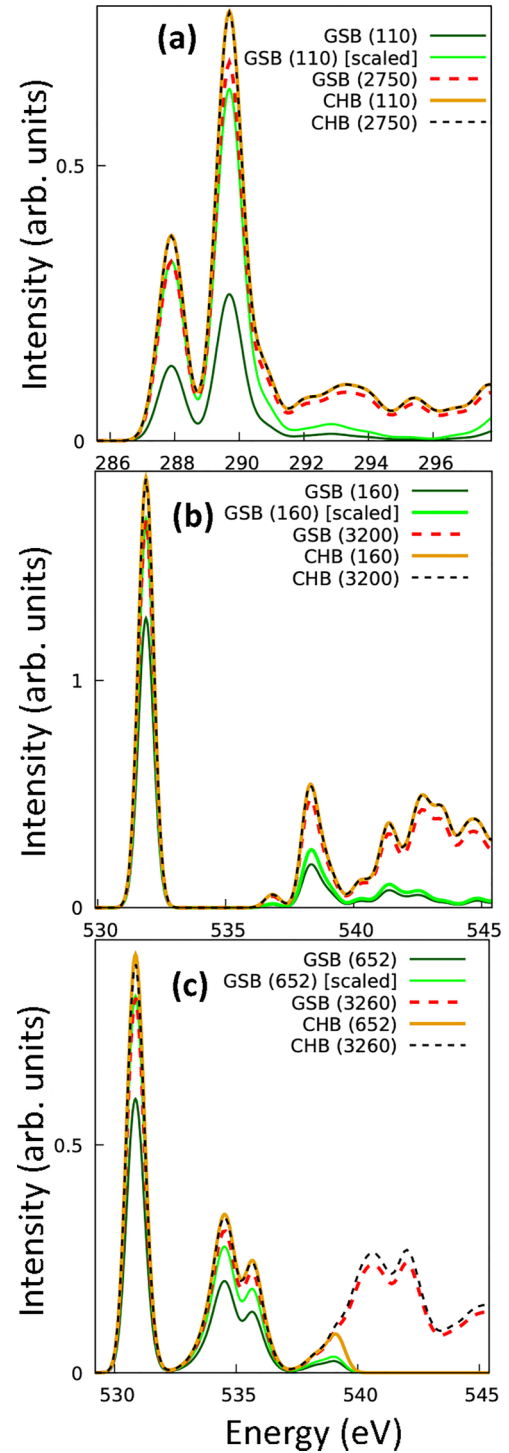


FIG. 1. (a) Methanol C  $K$ -edge, (b) acetone O  $K$ -edge, and (c) CuO O  $K$ -edge absorption spectra. The number in parentheses denotes the total number of FCH KS orbitals (excluding the core orbitals) used in the calculation. The “scaled” curve in each panel is obtained by multiplying the dark-green one by a constant factor in order to match the first-peak intensity with that of the broken red one.

for step 6 which now follows Eq. (11) and thereby involves a sum over ( $N + 1$ ) terms. Note that, in the CHB approach, step 4 is replaced by a calculation of the single-particle transition matrix elements  $\tilde{o}_l$  over *all* of the FCH-state orbitals,

TABLE I. Table showing for methanol C  $K$  edge, acetone O  $K$  edge, and CuO O  $K$  edge the estimates of core hours (parallelized across multiple 2.50 GHz Intel Xeon 20-core processors) used in the XAS calculations using the GSB and the CHB approaches. The number in parentheses denotes the total number of FCH KS orbitals (excluding the core orbitals) used in the calculation.

System	Core hours (GSB)	Core hours (CHB)
Methanol C $K$ (2750)	95	46
Acetone O $K$ (3200)	166	82
CuO O $K$ (3260)	1056	597

occupied and unoccupied alike, and therefore, should be a little more expensive than the corresponding GSB calculation. However, typically, for a calculation involving a large number of unoccupied orbitals, steps 2 and 3 are, by a large margin, the most expensive ones computationally. Therefore, the CHB approach, which, for the ground-state system, does not require step 2 at all and reduces the computational expense of step 3, is generally much cheaper than the GSB approach. In fact, assuming that, for a given system, steps 2 and 3 have comparable costs for the GS and the FCH state, the CHB approach is expected to reduce the total computational cost approximately by half. This, in fact, can be seen from Table I, which shows the estimates of the computational time required for methanol C  $K$ -edge, acetone O  $K$ -edge, and CuO O  $K$ -edge calculations with a relatively large number of unoccupied orbitals.

### B. Separation of spectral contributions

Note from Eq. (11) that the MBXAS amplitude  $A_{GS \rightarrow f}$  is essentially a linear combination, for  $l = 1, \dots, N$  and  $l = f$ , of single-particle transition amplitudes  $\tilde{o}_l$  from the core orbital to the orbital  $|\tilde{\phi}_l\rangle$ , with the factor  $\langle \Psi_{FCH}^{+f-l+0} | \Psi_{GS} \rangle$  serving as the coefficient. We can naturally separate these terms according to their physical origins as follows:

(1) The  $l = f$  term essentially corresponds to the expected quasiparticle transition, from the core orbital to orbital  $|\tilde{\phi}_f\rangle$ , which is formally unoccupied in  $|\Psi_{FCH}\rangle$ , but indirectly includes some perturbative effects of the presence of the core hole (since it is calculated based on the core-excited SCF).

(2) The  $l = 1, \dots, N$  terms accounts for the collective response of the rest of the electron density to the perturbation caused by the core hole by considering single-particle contributions from each of the occupied valence orbitals of  $|\Psi_{FCH}\rangle$ .

This separation is just a natural consequence of the many-body transition operator coupling the  $N$ -electron ground and core-excited states. However, another way to understand it is to consider the equivalence of absorption amplitude to the (stimulated) emission amplitude. We can rewrite Eq. (11) as

$$A_{GS \rightarrow f} = (-1)^\gamma \left[ \sum_{l=1, \dots, N, f} \langle \Psi_{FCH}^{+f} | \hat{O} | \Psi_{FCH}^{+f-l+0} \rangle \langle \Psi_{FCH}^{+f-l+0} | \Psi_{GS} \rangle \right]. \quad (26)$$

Now we can interpret this as a deexcitation from  $|\Psi_{FCH}^{+f}\rangle$  to  $|\Psi_{GS}\rangle$ , where the  $l$ th representative term can be recog-

nized as the amplitude  $\langle \Psi_{FCH}^{+f} | \hat{O} | \Psi_{FCH}^{+f-l+0} \rangle$  of a *non-self-consistent* deexcitation from the state  $|\Psi_{FCH}^{+f}\rangle$  to the core-filled state  $|\Psi_{FCH}^{+f-l+0}\rangle$ , multiplied, for the sake of restoring self-consistency, by the projection of this non-self-consistent state onto the actual final state  $|\Psi_{GS}\rangle$ . Thus, each of the  $(N + 1)$  terms in the sum corresponds to deexcitation of an electron from one of the occupied levels of  $|\Psi_{FCH}^{+f}\rangle$  to the core level.

### C. Comparison with the single-particle formalism

Armed with the numerically converged CHB treatment, we can now examine the relative importance of its various contributions. According to the expansion shown in Eq. (11),  $\tilde{o}_l$  is one among the  $N + 1$  terms ( $l = 1, \dots, N, f$ ) contributing to the transition amplitude. This affords a straightforward avenue for comparison of the MBXAS formalism with the more common single-particle formalism of XAS, in which the transition amplitude is approximated simply by  $\tilde{o}_f$ , and therefore, the absorption cross section is given by

$$\sigma(\omega) \propto \omega \sum_f |\langle \tilde{\phi}_f | \hat{O} | \tilde{\phi}_0 \rangle|^2 \delta(E_f - E_{GS} - \hbar\omega). \quad (27)$$

Mathematically, it can be shown that this truly single-particle approximation would be true if the valence occupied subspace of the ground and the FCH state were identical. Under such conditions,  $\langle \Psi_{FCH}^{+f-l+0} | \Psi_{GS} \rangle = \delta_{f,l}$ . Therefore, from Eq. (26), we have  $A_{GS \rightarrow f} = \tilde{o}_f$  (possibly with an inconsequential change in sign) and consequently the MBXAS cross section [Eq. (4)] reduces to the single-particle cross section [Eq. (27)]. We could now ask whether the degree to which the valence occupied subspaces differ is a useful metric to indicate a larger difference between  $A_{GS \rightarrow f}$  (i.e., the MBXAS transition amplitude) and  $\tilde{o}_f$ . We explore this possibility through a detailed analysis of the first peak of the C and O  $K$ -edge absorption spectra of an isolated methanol molecule.

As stated concisely in Eq. (11), the MBXAS amplitude is essentially a sum of the terms

$$C_l^{(f)} = (-1)^\gamma \tilde{o}_l \langle \Psi_{FCH}^{+f-l+0} | \Psi_{GS} \rangle \text{ for } l = 1, \dots, N, f, \quad (28)$$

for a given core-excited state  $|\Psi_{FCH}^{+f}\rangle$  labeled by core-excited (formerly unoccupied) FCH orbital index  $f$ . We will now assess which  $C_l^{(f)}$  are strong contributors to specific x-ray absorption transitions in methanol, by reference to Figs. 2 and 3.

At the C  $K$  edge in Fig. 2(a), the most intense XAS peak for photons polarized along the  $x$  axis (see Fig. 3 for the molecular orientation) occurs just below 290 eV and corresponds to transitions to  $|\Psi_{FCH}^{+11}\rangle$  for the core-excited C. As expected, the dominant contribution to this transition is  $C_{11}^{(11)}$ , defined by the  $f = 11$  orbital (FCH LUMO + 3) that labels this excited state. However, in addition, we see in Fig. 2(c) that there are also significant transition amplitude contributions from  $C_3^{(11)}$  and  $C_5^{(11)}$  which involve occupied orbitals 3 and 5 from this FCH SCF. As expected, each of these orbitals (3, 5, 11) displays noticeable  $p_x$  character on the core-excited C atom, as shown in Figs. 3(a)–3(c).

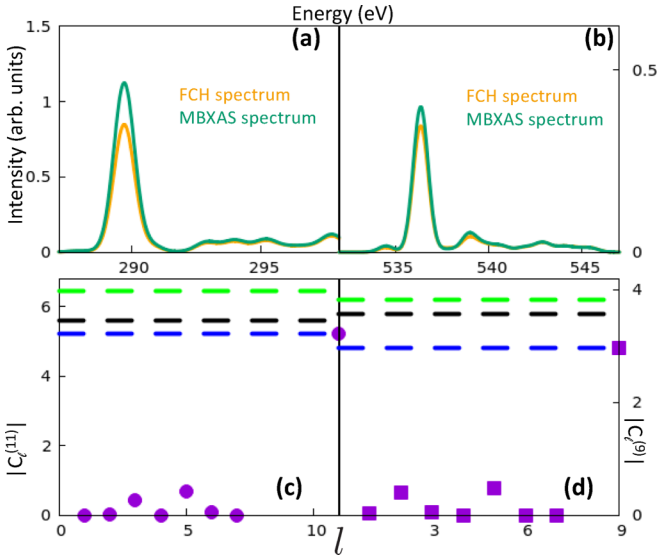


FIG. 2. Panels (a) and (b) show, for the methanol molecule, the C  $K$ -edge XAS spectrum for incident polarization along the  $x$  axis and the O  $K$ -edge XAS spectrum for incident polarization along the  $y$  axis, respectively (see molecular polarization in Fig. 3). Panels (c) and (d) analyze, in detail, the orbital contributions responsible for MBXAS transition amplitude of the most intense peaks seen in panels (a) and (b), respectively. Each panel contains discrete points showing the values of  $|C_l^{(f)}|$  as a function of  $l$  for a fixed  $f$  [ $f = 11$  in panel (c) and  $f = 9$  in panel (d)]. Since there are 7 valence electrons in the FCH state, as expected from Eq. (28), the discrete points correspond to  $l = 1, \dots, 7, 11$  in panel (c) and  $l = 1, \dots, 7, 9$  in panel (d). In panel (c) [panel (d)] the green, black, and blue lines correspond to  $|A_{GS \rightarrow f}|$ ,  $|\tilde{\delta}_f|$ , and  $|C_f^{(9)}|$  ( $|A_{GS \rightarrow f}|$ ,  $|\tilde{\delta}_f|$ , and  $|C_f^{(9)}|$ ), respectively. In other words, they denote the absolute amplitudes of the MBXAS transition, single-particle FCH transition, and the contribution of the  $l = f$  transition to MBXAS, respectively.

On the other hand, at the O  $K$  edge in Fig. 2(b), the most intense XAS peak for photons polarized along the  $y$  axis occurs above 536 eV and corresponds to transitions to  $|\Psi_{\text{FCH}}^{+9}\rangle$  for the core-excited O. As before, the dominant contribution to this transition is  $C_9^{(9)}$ , defined by the  $f = 9$  orbital (FCH LUMO + 1) that labels this excited state. However, in addition, we see in Fig. 2(d) that there are also significant transition amplitude contributions from  $C_2^{(9)}$  and  $C_5^{(9)}$  which involve occupied orbitals 2 and 5 from this FCH SCF. As expected, each of these orbitals (2, 5, 9) displays noticeable  $p_y$  character on the core-excited O atom, as shown in Figs. 3(d)–3(f).

The computed absolute values of the valence overlap between  $|\Psi_{\text{FCH}}\rangle$  and  $|\Psi_{\text{GS}}\rangle$  are 0.94 and 0.83 for the C and the O  $K$  edge of methanol, respectively. Since the excitation is expected to induce only negligible change in the tightly bound core orbital, the aforementioned valence overlap can be used as a good approximation for  $\langle \Psi_{\text{FCH}}^{+0} | \Psi_{\text{GS}} \rangle$ . As expected, owing to a higher overlap, the difference between  $|C_{11}^{(11)}|$  and  $|\tilde{\delta}_{11}|$  is smaller at the C  $K$  edge than that between  $|C_9^{(9)}|$  and  $|\tilde{\delta}_9|$  at the O  $K$  edge. However, the relative difference between the absolute values of the MBXAS and the single-particle amplitude, i.e.,  $\frac{|A_{GS \rightarrow f}| - |\tilde{\delta}_f|}{|\tilde{\delta}_f|}$  [this is the relative difference between the green and black horizontal lines in Figs. 2(c) and 2(d)] is

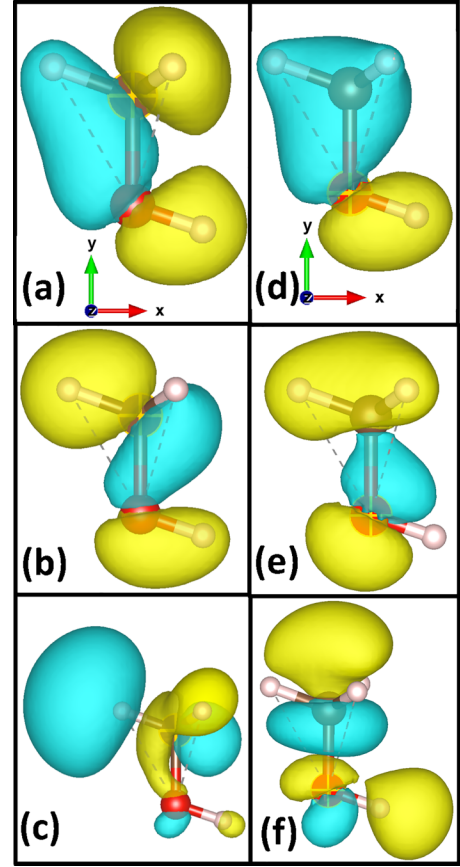


FIG. 3. Isovalue plots of occupied KS orbitals of the FCH states of the methanol molecule resulting from core ionization of specific atoms. The red, brown, and white spheres denote the oxygen, carbon, and hydrogen atoms, respectively. The core-ionized atom is marked by a “+” sign. The strongly contributing orbitals to the XAS intensity are indexed according to the energy ordering of their KS eigenvalues as (a) 3, (b) 5, (c) 11 (LUMO + 3) for the C  $K$  edge and (d) 2, (e) 5, (f) 9 (LUMO + 1) for the O  $K$  edge.

much higher for C (0.15) than it is for O (0.07), demonstrating that a lower overlap between the valence and the FCH occupied subspaces does not necessarily imply a larger difference between  $|A_{GS \rightarrow f}|$  and  $|\tilde{\delta}_f|$ . In general, the formally more accurate option is clearly to include all additional valence occupied contributions, although it is clear that the XAS is somewhat dominated by its single-particle contributions in this case. We will explore a counterexample below.

#### D. Auxiliary orbital

It is important to note that the excitation described by MBXAS is an inherently many-body process with no physical single-particle analog. However, the transition amplitude given by Eq. (11) inspires the formulation of a hypothetical single-particle excitation of a core electron to an auxiliary orbital given by

$$|\tilde{\phi}_f^{\text{aux}}\rangle = \sum_{l=1, \dots, N, f} (-1)^{\nu_l} \langle \Psi_{\text{GS}} | \Psi_{\text{FCH}}^{+f-l+0} \rangle |\tilde{\phi}_l\rangle, \quad (29)$$

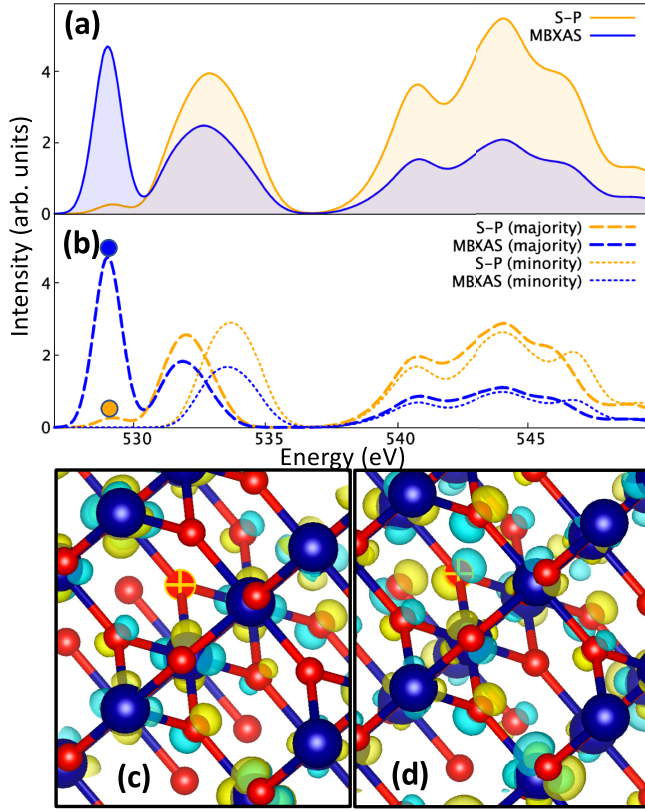


FIG. 4. Panel (a) shows the O  $K$ -edge XAS of  $\text{CrO}_2$  simulated using the single-particle (S-P) and MBXAS methods. Panel (b) shows separately the majority and minority spin contributions to spectra of both methods. The small orange (blue) circle corresponds to the first peak in the majority-spin spectrum and is analyzed using the isovalue plot of panel (c) [panel (d)] which shows, at the  $\Gamma$  point, the orbital  $\tilde{\phi}_f(\mathbf{r})$  [ $\tilde{\phi}_f^{\text{aux}}(\mathbf{r})$ ] that contributes most strongly to this peak. In this case,  $f$  corresponds to the lowest unoccupied level of the FCH state. The core-ionized oxygen atom is marked by a “+” symbol in both lower panels. The blue and the red spheres correspond to the Cr and the O atoms, respectively.

such that the amplitude of this excitation equals that of the actual many-body one, i.e.,

$$\tilde{\sigma}_f^{\text{aux}} = \langle \tilde{\phi}_f^{\text{aux}} | \hat{\sigma} | \phi_0 \rangle = A_{\text{GS} \rightarrow f}, \quad (30)$$

as can be seen by comparison with Eq. (9). The utility of the auxiliary orbital can be more easily appreciated by looking at the O  $K$ -edge spectrum of the ferromagnetic, crystalline solid  $m - \text{CrO}_2$  which was discussed in detail in Ref. [36].

As shown in Fig. 4(b), for the majority spin channel of  $\text{CrO}_2$ , the intensity of the O  $K$ -edge first peak at 529 eV is severely underestimated by the single-particle approximation. This underestimation can be understood by examining, at the  $\Gamma$  point, the corresponding excited single-particle orbital in the isovalue plots. The filled FCH orbital  $f$ , in this case, sits at the conduction band minimum (CBM), or first unoccupied orbital (analogous to LUMO at the  $\Gamma$  point) of the FCH state. We see from Fig. 4(c) that  $\tilde{\phi}_{\text{LUMO}}(\mathbf{r})$  has apparently no visible presence on the core-ionized O atom, which explains the almost vanishing intensity of this first peak in the single-particle spectrum. On the other hand, the associated auxiliary

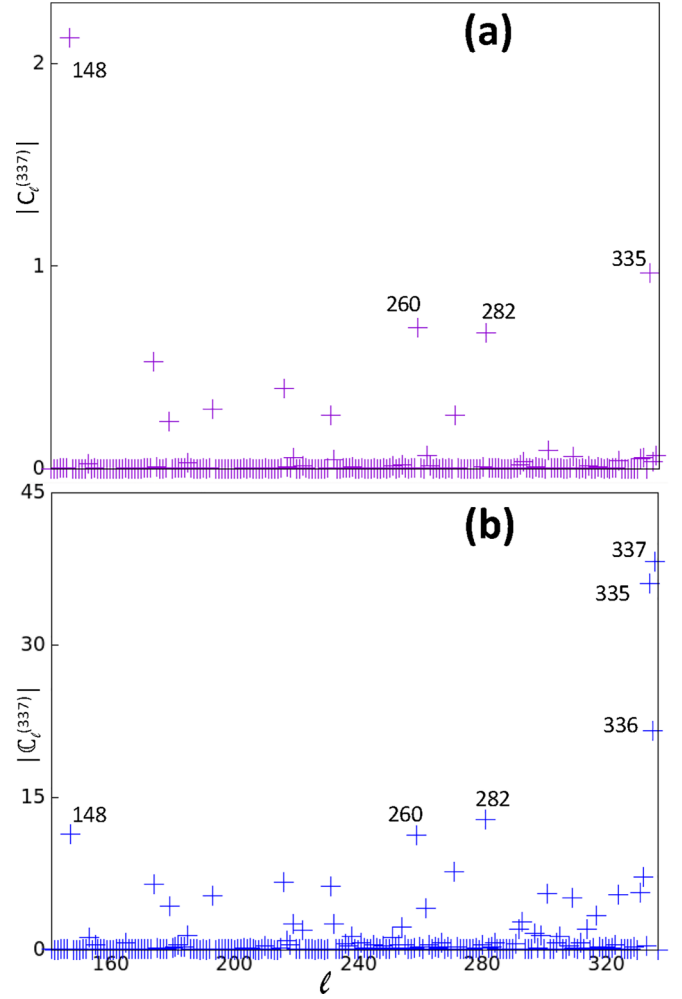


FIG. 5. For O  $K$ -edge electronic transition at the  $\Gamma$  point of  $\text{CrO}_2$ , panels (a) and (b) show, as a function of the FCH KS orbital indexed by  $l$ , the absolute values  $|C_l^{(337)}|$  [see Eq. (28)] and  $|C_l^{(337)}|$ , respectively, where the 337th FCH orbital is the lowest unoccupied one.

orbital  $\tilde{\phi}_{\text{LUMO}}^{\text{aux}}(\mathbf{r})$  shown in Fig. 4(d) exhibits substantial O- $p$  character on the core-ionized atom, due to inclusion of contributions from occupied valence orbitals in Eq. (29), and hence gives rise to a drastically higher (and more accurate) MBXAS intensity.

In a nutshell, the concept of auxiliary orbitals is introduced to provide a visual representation of the excited electron (similar, in spirit, to the natural transition orbitals [42–45] used in quantum chemistry) that is in agreement with the MBXAS intensity. The definition of an auxiliary orbital  $\tilde{\phi}_f^{\text{aux}}$  incorporates the relevant contributions from all the FCH occupied orbitals in such a way that the single-particle transition probability of promoting the core-electron to the orbital  $\tilde{\phi}_f^{\text{aux}}$  equals the true (i.e., MBXAS) transition probability of exciting the core electron to the KS orbital  $\tilde{\phi}_f$ .

Figure 5(a) shows, as a function of the FCH orbital  $\tilde{\phi}_l$ , the absolute value of  $C_l^{(\text{LUMO})}$  [see Eq. (28)], which is essentially the contribution to the first-peak MBXAS amplitude [Eq. (11)] of the single-particle transition associated with the orbital  $\tilde{\phi}_l$ . On the other hand, Fig. 5(b) shows the



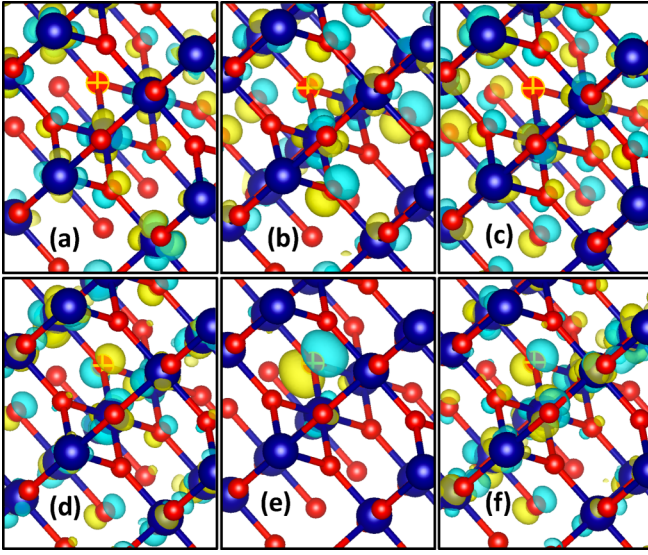


FIG. 6. Isovalue plots of  $\Gamma$ -point FCH KS orbitals contributing most strongly to the auxiliary orbital  $\tilde{\phi}_f^{\text{LUMO}}(\mathbf{r})$  shown in Fig. 4(d). Panels (a) through (f) indicate decreasing contributions  $|\mathcal{C}_l^{\text{LUMO}}|$  starting from the largest: (a) orbital 337 (LUMO), (b) 335 (HOMO-1), (c) 336 (HOMO), (d) 282, (e) 148, and (f) 260. The corresponding contributions can be found in Fig. 5(b).

absolute value of  $\mathcal{C}_l^{\text{(LUMO)}} = (-1)^{\gamma_l} \langle \Psi_{\text{GS}} | \Psi_{\text{FCH}}^{+\text{LUMO}-l+0} \rangle$ , which, as seen from Eq. (29), serves as the contribution of  $\tilde{\phi}_l$  to the auxiliary orbital  $\tilde{\phi}_{\text{LUMO}}^{\text{aux}}$ . As expected, due to the factor of  $\tilde{\delta}_l$ , these two contributions do not follow the same trend. Panel (a) of Fig. 5 clearly shows that the largest contribution to the MBXAS amplitude arises from the 148th FCH orbital  $\tilde{\phi}_{148}$ , situated deep in the valence subspace. Predictably,  $\tilde{\phi}_{148}$  exhibits strong  $p$  character on the excited oxygen atom, as seen from panel (e) of Fig. 6, which shows the isovalue plots of FCH orbitals with appreciable contribution in  $\tilde{\phi}_{\text{LUMO}}^{\text{aux}}$ . Conversely, as evident from Fig. 5(b), the largest contribution to the auxiliary orbital comes from the FCH LUMO, which, as already seen from Fig. 4(c), has negligible  $p$  character on the core-excited oxygen atom. For the auxiliary orbital  $\tilde{\phi}_{\text{LUMO}}^{\text{aux}}$ , the noticeable  $p$  character on the excited atom can then be attributed to the cumulative contributions from the O- $p$  character of various valence orbitals having appreciable  $\mathcal{C}_l^{\text{(LUMO)}}$ .

Note, from Eq. (23), that auxiliary orbitals can also be defined in the GSB approach as

$$|\tilde{\phi}_{\text{aux}}^f\rangle = \sum_{c>N} \langle \Psi_{\text{FCH}}^{+f} | \Psi_{\text{GS}}^{+c-0} \rangle |\phi_c\rangle. \quad (31)$$

However, due to the convergence issues discussed above, this might require the evaluation and manipulation of a large number of GS unoccupied orbitals, rendering the calculation formidably expensive.

### E. The $f^{(n)}$ contribution

Up to this point, each final state has been modeled by adding an electron to an unoccupied orbital of the FCH state so that each final state contains a single hole (i.e., the core hole) and a single excited electron (the one added on top of

the FCH state). At this point we would like to briefly touch upon the contribution of the so-called  $f^{(n)}$  terms [36] within the MBXAS formalism, whereby the neutral final state is generated by exciting, in addition to the aforementioned e-h pair consisting of the core hole and the conduction electron, additional e-h pairs outside the core region (i.e., pairs of valence hole and conduction electron). In other words, a final state  $|\Psi_f^{(n)}\rangle$  corresponding to the  $f^{(n)}$  term contains, in total,  $n$  number of holes (a single core hole and  $n-1$  number of valence holes) and  $n$  number of conduction electrons. Such an excited state is approximated as

$$\begin{aligned} |\Psi_f^{(n)}\rangle &= |\Psi_{\text{FCH}}^{+f_{n-1}+\dots+f_1+f-h_{n-1}-\dots-h_1}\rangle \\ &= (-1)^\Lambda \tilde{a}_{f_{n-1}}^\dagger \dots \tilde{a}_{f_1}^\dagger \tilde{a}_f^\dagger \tilde{a}_{h_{n-1}} \dots \tilde{a}_{h_1} |\Psi_{\text{FCH}}\rangle, \end{aligned} \quad (32)$$

with  $\Lambda = (nN - h_1 - \dots - h_{n-1})$ , obtained by generating, in addition to the core excitation,  $n-1$  number of electron-hole pairs outside the core region, with the  $m$ th conduction electron (valence hole) being indexed by  $f_m$  ( $h_m$ ). Without loss of generality, the orbital indices are made to follow the order  $f_{n-1} > \dots > f_1 > f > h_{n-1} > \dots > h_1$ . The core-excited state shown in Eq. (3), which contains no valence hole, can immediately be recognized as  $|\Psi_f^{(1)}\rangle$ .

Within the CHB approach, the transition amplitude associated with  $|\Psi_f^{(n)}\rangle$  can then be expressed as

$$\begin{aligned} A_{\text{GS}\rightarrow\text{f}}^{(n)} &= \langle \Psi_f^{(n)} | \sum_l \tilde{\delta}_l \tilde{a}_l^\dagger \tilde{a}_0 | \Psi_{\text{GS}} \rangle \\ &= \sum_l (-1)^{\gamma_l^{(n)}} \tilde{\delta}_l \langle \Psi_{\text{FCH}}^{+f_{n-1}+\dots+f_1+f-h_{n-1}-\dots-h_1-l+0} | \Psi_{\text{GS}} \rangle, \end{aligned} \quad (33)$$

where

$$|\Psi_{\text{FCH}}^{+f_{n-1}+\dots+f_1+f-h_{n-1}-\dots-h_1-l+0}\rangle = (-1)^{\gamma_l^{(n)}} \tilde{a}_l \tilde{a}_0^\dagger |\Psi_f^{(n)}\rangle, \quad (34)$$

with the term  $\gamma_l^{(n)}$  depending on the relative position of  $l$  with respect to  $h_1 \dots h_{n-1}$ . Equations (33) and (34) reveal that, within the CHB approach, the GS unoccupied orbitals are not needed to compute the transition amplitude  $A_{\text{GS}\rightarrow\text{f}}^{(n)}$  for any value of  $n$ . Similar to the special case of  $f^{(1)}$  [see Eq. (29)], a general transition arising from the  $f^{(n)}$  term can be associated with the auxiliary orbital

$$\begin{aligned} |\tilde{\phi}_{\text{aux}}^{f^{(n)}}\rangle &= \sum_l (-1)^{\gamma_l^{(n)}} \langle \Psi_{\text{GS}} | \Psi_{\text{FCH}}^{+f_{n-1}+\dots+f_1+f-h_{n-1}-\dots-h_1-l+0} \rangle |\tilde{\phi}_l\rangle. \end{aligned} \quad (35)$$

As originally defined, we reiterate that within the MBXAS formalism, the contribution of  $f^{(n)}$  for  $n > 1$  ignores interactions between the subsequent electron-hole pairs and, therefore, is really only accurate for studies of metallic systems. It is especially inaccurate for insulating/semiconducting systems which should exhibit significant exciton binding for additional valence electron-hole pairs.

#### IV. CONCLUSION

In summary, in this article, we have presented an efficient representation of the determinant-based many-body formalism for simulating x-ray absorption spectroscopy, called MBXAS. We have shown that by recasting the relevant equations in terms of the CHB orbitals, one can circumvent the need for calculating the unoccupied orbitals of the ground state, leading to appreciable reduction in the computational expenses. Additionally, the reformulated expression for the transition amplitude does not suffer from convergence issues since it is now limited to a sum of a finite number of terms. Within the CHB approach, the amplitude is expanded as a linear combination of independent-electron transitions from various core-excited state orbitals. Therefore, such a treatment facilitates a direct comparison of the MBXAS formalism with the so-called single-particle approximation. As we show analytically and with a representative example, the aforementioned comparison motivates the association of a given MBXAS transition to a purely single-particle electronic excitation from the core level to a hypothetical auxiliary orbital. A visual representation of this auxiliary orbital offers intuitive justification for the relevant absorption intensity.

#### ACKNOWLEDGMENTS

This work was carried out at The Molecular Foundry (TMF), Lawrence Berkeley National Laboratory (LBNL). Financial support for this work was provided by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Computational works were performed using the supercomputing resources of the National Energy Research Scientific Computing Center (NERSC) and the TMF clusters managed by the High Performance Computing Services Group, at LBNL.

#### APPENDIX A: DERIVATION OF EQ. (10)

While constructing a many-electron state from the empty state  $|\emptyset\rangle$  by applying a set of creation operators, let us follow the following convention for the operator ordering: if the energy of the  $i$ th orbital is lower than that of the  $j$ th orbital, then the  $i$ th creation operator will be placed to the right (not necessarily immediate right) of the  $j$ th creation operator. Then, the state  $|\Psi_f\rangle$  [see Eq. (3)] can be written as

$$|\Psi_f\rangle = \tilde{a}_f^\dagger \tilde{a}_N^\dagger \dots \tilde{a}_{l+1}^\dagger \tilde{a}_l^\dagger \tilde{a}_{l-1}^\dagger \dots \tilde{a}_1^\dagger |\emptyset\rangle. \quad (\text{A1})$$

This gives us

$$\begin{aligned} \tilde{a}_0^\dagger \tilde{a}_l |\Psi_f\rangle &= \tilde{a}_0^\dagger \tilde{a}_l \tilde{a}_f^\dagger \tilde{a}_N^\dagger \dots \tilde{a}_{l+1}^\dagger \tilde{a}_l^\dagger \tilde{a}_{l-1}^\dagger \dots \tilde{a}_1^\dagger |\emptyset\rangle \\ &= (-1)^{(N-l+1)} \tilde{a}_0^\dagger \tilde{a}_f^\dagger \tilde{a}_N^\dagger \dots \tilde{a}_{l+1}^\dagger \tilde{a}_l \tilde{a}_l^\dagger \tilde{a}_{l-1}^\dagger \dots \tilde{a}_1^\dagger |\emptyset\rangle \\ &= (-1)^{(N-l+1)} \tilde{a}_0^\dagger \tilde{a}_f^\dagger \tilde{a}_N^\dagger \dots \tilde{a}_{l+1}^\dagger \tilde{a}_{l-1}^\dagger \dots \tilde{a}_1^\dagger |\emptyset\rangle \\ &= (-1)^{(N-l+1)} \cdot (-1)^N \tilde{a}_f^\dagger \tilde{a}_N^\dagger \dots \tilde{a}_{l+1}^\dagger \tilde{a}_{l-1}^\dagger \dots \tilde{a}_1^\dagger \tilde{a}_0^\dagger |\emptyset\rangle. \end{aligned} \quad (\text{A2})$$

The second line is obtained by moving the operator  $\tilde{a}_l$  to the right in order to place it to the left of the operator  $\tilde{a}_l^\dagger$ . For this,  $\tilde{a}_l$  was moved by  $N+1-l$  places since Eq. (A1) contains  $(N+1)$  operators in total. The third line is obtained from the second simply by omitting  $\tilde{a}_l$  and  $\tilde{a}_l^\dagger$ . In the fourth line,  $\tilde{a}_0^\dagger$  was moved by  $N$  places to the extreme right, immediately next to  $|\emptyset\rangle$ . Now, if  $|\Psi_{\text{FCH}}^{+f-l+0}\rangle$  denotes a core-excited state obtained from  $|\Psi_{\text{FCH}}\rangle$  *non-self-consistently* by adding an electron to the unoccupied orbital  $f$  and removing an electron from the occupied orbital  $l$ , then, following the aforementioned convention for operator ordering we can appreciate the equality

$$|\Psi_{\text{FCH}}^{+f-l+0}\rangle = \tilde{a}_f^\dagger \tilde{a}_N^\dagger \dots \tilde{a}_{l+1}^\dagger \tilde{a}_{l-1}^\dagger \dots \tilde{a}_1^\dagger \tilde{a}_0^\dagger |\emptyset\rangle. \quad (\text{A3})$$

Then, plugging Eq. (A3) into Eq. (A2), we recover Eq. (10).

#### APPENDIX B: COMPUTATIONAL DETAILS

All DFT calculations, including structural relaxation of the systems, are performed with plane-wave basis functions and the PBE exchange-correlation functional implemented in the Quantum Espresso package. For the calculations on methanol, acetone, and CuO, ultrasoft pseudopotentials have been used (Fig. 1 requires the computation of a large number of unoccupied KS orbitals; this would be extremely expensive with norm-conserving pseudopotentials) with energy cutoffs of 30 eV and 240 eV for wave function and density, respectively. Due to technical reasons related to the Quantum Espresso package, calculating a linear combination of multiple complex orbitals, which, as per Eq. (29), is required for obtaining the auxiliary orbital, is complicated with ultrasoft pseudopotentials. This is why, for the calculation on CrO<sub>2</sub>, norm-conserving pseudopotentials have been used with energy cutoffs of 90 eV and 360 eV for wave function and density, respectively. The single-particle wave functions are expressed in terms of a set of optimal basis functions [46,47].

[1] V. Coropceanu, H. Li, P. Winget, L. Zhu, and J.-L. Brédas, *Annu. Rev. Mater. Res.* **43**, 63 (2013).  
 [2] R. P. Messmer, *J. Non-Cryst. Solids* **75**, 285 (1985).  
 [3] P. J. Hasnip, K. Refson, M. I. J. Probert, J. R. Yates, S. J. Clark, and C. J. Pickard, *Philos. Trans. R. Soc. A* **372**, 20130270 (2014).  
 [4] A. Y. S. Eng, C. B. Soni, Y. Lum, E. Khoo, Z. Yao, S. K. Vineeth, V. Kumar, J. Lu, C. S. Johnson, C. Wolverton, and Z. W. Seh, *Sci. Adv.* **8**, eabm2422 (2022).

[5] Q. He, B. Yu, Z. Li, and Y. Zhao, *Energy Environ. Mater.* **2**, 264 (2019).  
 [6] F. G. Eich, M. Di Ventura, and G. Vignale, *Phys. Rev. Lett.* **112**, 196401 (2014).  
 [7] J. Shi and S. Yun, in *Counter Electrodes for Dye-Sensitized and Perovskite Solar Cells* (John Wiley & Sons, Weinheim, Germany, 2018), Chap. 19, pp. 487–509.  
 [8] N. Schaduangrat, S. Lampa, S. Simeon, M. P. Gleeson, O. Spjuth, and C. Nantasenamat, *J. Cheminf.* **12**, 9 (2020).

- [9] O. A. Arodola and M. E. Soliman, *Drug Des. Dev. Ther.* **11**, 2551 (2017).
- [10] N. Ye, Z. Yang, and Y. Liu, *Drug Discovery Today* **27**, 1411 (2022).
- [11] S. Sanvito, C. Oses, J. Xue, A. Tiwari, M. Zic, T. Archer, P. Tozman, M. Venkatesan, M. Coey, and S. Curtarolo, *Sci. Adv.* **3**, e1602241 (2017).
- [12] F. de Groot, *Chem. Rev.* **101**, 1779 (2001).
- [13] J. W. Smith and R. J. Saykally, *Chem. Rev.* **117**, 13909 (2017).
- [14] D. Prendergast and G. Galli, *Phys. Rev. Lett.* **96**, 215502 (2006).
- [15] Q. Li, R. Qiao, L. A. Wray, J. Chen, Z. Zhuo, Y. Chen, S. Yan, F. Pan, Z. Hussain, and W. Yang, *J. Phys. D* **49**, 413003 (2016).
- [16] M. Freiwald, S. Cramm, W. Eberhardt, and S. Eisebitt, *J. Electron Spectrosc. Relat. Phenom.* **137-140**, 413 (2004).
- [17] X. Liu and T.-C. Weng, *MRS Bull.* **41**, 466 (2016).
- [18] T. A. Pascal, C. D. Pemmaraju, and D. Prendergast, *Phys. Chem. Chem. Phys.* **17**, 7743 (2015).
- [19] L. F. Wan, J. Wright, B. R. Perdue, T. T. Fister, S. Kim, C. A. Appleby, and D. Prendergast, *Phys. Chem. Chem. Phys.* **18**, 17326 (2016).
- [20] R. Qiao, I. T. Lucas, A. Karim, J. Syzdek, X. Liu, W. Chen, K. Persson, R. Kostecki, and W. Yang, *Adv. Mater. Interfaces* **1**, 1300115 (2014).
- [21] S. Roychoudhury, R. Qiao, Z. Zhuo, Q. Li, Y. Lyu, J.-H. Kim, J. Liu, E. Lee, B. J. Polzin, J. Guo, S. Yan, Y. Hu, H. Li, D. Prendergast, and W. Yang, *Energy Environ. Mater.* **4**, 246 (2021).
- [22] Y. Horikawa, H. Arai, T. Tokushima, and S. Shin, *Chem. Phys. Lett.* **522**, 33 (2012).
- [23] R. Qiao, L. A. Wray, J.-H. Kim, N. P. W. Pieczonka, S. J. Harris, and W. Yang, *J. Phys. Chem. C* **119**, 27228 (2015).
- [24] C. Branci, M. Womes, P. Lippens, J. Olivier-Fourcade, and J. Jumas, *J. Solid State Chem.* **150**, 363 (2000).
- [25] S. Roychoudhury, Z. Zhuo, R. Qiao, L. Wan, Y. Liang, F. Pan, Y.-d. Chuang, D. Prendergast, and W. Yang, *ACS Appl. Mater. Interfaces* **13**, 45488 (2021).
- [26] N. A. Besley and F. A. Asmuruf, *Phys. Chem. Chem. Phys.* **12**, 12024 (2010).
- [27] G. Fronzoni, R. De Francesco, and M. Stener, *J. Phys. Chem. B* **109**, 10332 (2005).
- [28] D. Hait, K. J. Oosterbaan, K. Carter-Fenk, and M. Head-Gordon, *J. Chem. Phys.* **156**, 201104 (2022).
- [29] K. Lopata, B. E. Van Kuiken, M. Khalil, and N. Govind, *J. Chem. Theory Comput.* **8**, 3284 (2012).
- [30] E. L. Shirley, *Phys. Rev. Lett.* **80**, 794 (1998).
- [31] R. Laskowski and P. Blaha, *Phys. Rev. B* **82**, 205104 (2010).
- [32] W. Olovsson, I. Tanaka, T. Mizoguchi, P. Puschnig, and C. Ambrosch-Draxl, *Phys. Rev. B* **79**, 041102(R) (2009).
- [33] S. Roychoudhury, L. A. Cunha, M. Head-Gordon, and D. Prendergast, *Phys. Rev. B* **106**, 075133 (2022).
- [34] M. Taillefumier, D. Cabaret, A.-M. Flank, and F. Mauri, *Phys. Rev. B* **66**, 195107 (2002).
- [35] Y. Liang, J. Vinson, S. Pemmaraju, W. S. Drisdell, E. L. Shirley, and D. Prendergast, *Phys. Rev. Lett.* **118**, 096402 (2017).
- [36] Y. Liang and D. Prendergast, *Phys. Rev. B* **97**, 205127 (2018).
- [37] P. Jiang, D. Prendergast, F. Borondics, S. Porsgaard, L. Giovanetti, E. Pach, J. Newberg, H. Bluhm, F. Besenbacher, and M. Salmeron, *J. Chem. Phys.* **138**, 024704 (2013).
- [38] A. H. England, A. M. Duffin, C. P. Schwartz, J. S. Uejio, D. Prendergast, and R. J. Saykally, *Chem. Phys. Lett.* **514**, 187 (2011).
- [39] F. De Groot and A. Kotani, *Core Level Spectroscopy of Solids*, Advances in Condensed Matter Science (Taylor & Francis Group, Boca Raton, Florida, 2008).
- [40] Formally, the determinant is a signed magnitude of the exterior product of the  $N$ -dimensional basis, which has dimension 1.
- [41] L. Forestell and F. Marsiglio, *Can. J. Phys.* **93**, 1009 (2015).
- [42] A. V. Luzanov, A. A. Sukhorukov, and V. É. Umanskii, *Theor. Exp. Chem.* **10**, 354 (1976).
- [43] R. L. Martin, *J. Chem. Phys.* **118**, 4775 (2003).
- [44] M. Head-Gordon, A. M. Grana, D. Maurice, and C. A. White, *J. Phys. Chem.* **99**, 14261 (1995).
- [45] A. I. Krylov, *J. Chem. Phys.* **153**, 080901 (2020).
- [46] D. Prendergast and S. G. Louie, *Phys. Rev. B* **80**, 235126 (2009).
- [47] E. L. Shirley, *Phys. Rev. B* **54**, 16464 (1996).