

Real-space representation of the quasiparticle self-consistent GW self-energy and its application to defect calculations

Ozan Dernek , Dmitry Skachkov ,* and Walter R. L. Lambrecht 

Department of Physics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-7079, USA

Mark van Schilfgaarde 

National Renewable Energy Laboratory, Golden, Colorado 80401, USA



(Received 18 February 2022; revised 9 May 2022; accepted 13 May 2022; published 26 May 2022)

The quasiparticle self-consistent (QS) GW (G for Green's function, W for screened Coulomb interaction) approach incorporates the corrections of the quasiparticle energies from their Kohn-Sham density functional theory (DFT) eigenvalues by means of an energy-independent and Hermitian self-energy matrix usually given in the basis set of the DFT eigenstates. By expanding these into an atom-centered basis set (specifically here the linearized muffin-tin orbitals) a real space representation of the self-energy corrections becomes possible. We show that this representation is relatively short-ranged. This offers opportunities to construct the self-energy of a complex system from parts of the system by a cut-and-paste method. Specifically for a point defect, represented in a large supercell, the self-energy can be constructed from those of the host and a smaller defect-containing cell. The self-energy of the periodic host can be constructed simply from a GW calculation for the primitive cell. We show for the case of the As_{Ga} in GaAs that the defect part can already be well represented by a minimal eight-atom cell and allows us to construct the self-energy for a 64-atom cell in good agreement with direct QSGW calculations for the large cell. Using this approach to an even larger 216-atom cell shows the defect band approaches an isolated defect level. The calculations also allow us to identify a second defect band which appears as a resonance near the conduction band minimum. The results on the extracted defect levels agree well with Green's function calculations for an isolated defect and with experimental data.

DOI: [10.1103/PhysRevB.105.205136](https://doi.org/10.1103/PhysRevB.105.205136)

I. INTRODUCTION

In recent years, many-body-perturbation theory, specifically in Hedin's GW approximation [1,2], where G is the one-electron Green's function and W the screened Coulomb interaction, has emerged as the method of choice to calculate meaningful quasiparticle excitation energies as opposed to Kohn-Sham density-functional one-electron energies. In particular, this yields more accurate band structures and band gaps in good agreement with experiments, typically within ~ 0.1 eV depending somewhat on the details of the implementation and the material [3].

For localized levels, such as point defects, on the other hand, excitations are usually calculated by means of a ΔSCF approach from differences between two total energies. More precisely for point defects, it is now standard procedure to calculate the energies of formation as a function of Fermi level position (i.e., the electron chemical potential) in the gap and then to determine the transition energies, which are the crossing points of these energies of formation from one charge state to the other [4]. On the other hand, it has been pointed out that the GW quasiparticle energies for a defect system can be directly related to vertical excitations,

meaning excitations which keep the structure unchanged. Apart from excitonic effects, their differences represent the optical transitions, for example, for transferring an electron from a defect level to the conduction band or from the valence band to a defect level [5,6]. For example, transferring an electron from a defect level to the conduction band minimum (CBM) changes the defect from one charge state q to another $q + 1$, with an electron in a delocalized conduction band state. Thus, the difference in CBM and defect-level quasiparticle energies calculated at the fixed geometry of the q state is the vertical excitation energy for this process. Subsequently, one may add relaxation energies of the defect to lowest energy geometry within a given charge state [5]. Thus, the thermodynamic transition state level of the $q/q + 1$ transition may be obtained from a total energy relaxation within a given charge state calculated at the density functional theory (DFT) level combined with the quasiparticle excitation energies calculated at the GW level. These types of GW defect calculations were recently reviewed by Chen *et al.* [7,8]. Apart from the excitonic effect, the vertical transition is often directly of interest as an approximation to the optical transition.

In practice, defect levels are usually calculated in supercells using periodic boundary conditions, and the defect levels turn into defect bands. So, the above relation gives a renewed incentive to take defect one-electron band structures seriously, rather than dismissing them as irrelevant Kohn-Sham

*Current address: Department of Physics, University of Florida, Gainesville, Florida 32611, USA.

eigenvalues, provided they are calculated at the *GW* quasiparticle level.

Nonetheless, the *GW* method has not yet found widespread applications in defect studies. This is at least in part due to the large computational effort required for *GW* calculations. In particular, the latter is challenging for the large supercells needed to represent defects adequately. Thus, there is a need for improving the efficiency of the *GW* approach, eventually at the cost of some simplification, to make it applicable to larger systems.

On the other hand, in many defect calculations, the infamous gap underestimate of the typical semilocal generalized gradient approximation or local density approximations (LDAs) can lead to serious errors in defect calculations. Defect levels which should be in the gap may end up in the band continuum. This also affects the total energies and therefore transition levels if one considers charge states in which that defect level (now a resonance in the band) is given an extra charge because in the calculation that charge is actually placed in a delocalized state at the bottom of the band rather than in the defect level. Thus, there seem to be some advantages to start from a more accurate quasiparticle one-electron theory such as *GW*. Total energies within the QSGW approach are in principle calculable at the random phase approximation (RPA) level by means of an adiabatic connection approach [9] but are difficult to converge because they require a sum over unoccupied bands.

It would seem desirable to at least build into the calculation of the defect, the correct band structure of the host, overcoming thereby the band gap problem. Several such approaches have been used in the past: for example, Christensen [10] advocated using a δ -function corrected potential in which a δ function placed at the cation site and some interstitial sites, raises the *s*-like partial waves at those sites in energy and since these *s*-like orbitals form a predominant part of the conduction band, they artificially corrected the gap. LDA + *U* with several orbital-dependent *U* parameters were also shown to provide an effective way to mimic the corrected band structure [11–13]. Nonlocal external potentials (NLEPs) of the form $\Delta V_{\alpha,i}^{\text{NLEP}}$ adjusted to reproduce the conduction band structure were introduced by Lany *et al.* [14]. Modified pseudopotentials were also used for this purpose [15].

The question thus arises to what extent we can decouple the host band structure effects from the defect. Our goal with this project was to explore whether the *GW* self-energy of a defect system could be constructed from that of the host and the defect site itself or its immediate neighborhood without having to carry out the expensive full *GW* calculation for the large unit cell required to adequately represent a defect.

The most prevalent approach nowadays to incorporate the gap corrections beyond semi-local functions is to use a hybrid functional such as the Heyd-Scuseria-Ernzerhof [16–18]. That approach significantly improves the gaps by including a fraction of the exact exchange operator cutoff, usually beyond a certain range. By adjusting the fraction of the exchange included, the gap can be adjusted. While this approach correctly incorporates the gap correction, it is less clear that it also adjusts both band edges individually and/or obeys the generalized Koopmans theorem [19,20] for different defects simultaneously with the host band structure. Furthermore, it is

also a relatively expensive approach with computational effort well beyond that of a semilocal calculation. The approach we present here to construct the *GW* self-energy is also applicable to the nonlocal exact exchange and, hence, could also make that approach more efficient.

In this paper, we show that the QSGW self-energy matrix can be expanded in atom-centered orbitals, such as linearized muffin-tin orbitals (LMTOs). If these are chosen sufficiently localized, then the self-energy matrix can be represented in real space within a finite range. One might envision doing this also with maximally localized Wannier functions. This then offers opportunities to approximately construct the self-energy of a system by partitioning the system in subparts and constructing the self-energy by a cut-and-paste approach. In particular, we apply this here to point defects. We first construct the self-energy of the host in a supercell from that of the primitive cell. In a second step, we replace the part of the self-energy matrix related to the defect atom and its near neighborhood in terms of the self-energy of a smaller supercell containing the defect for which a *GW* calculation is more readily feasible. We validate the accuracy of the approach with the well-studied case of the As_{Ga} defect.

One might view our approach also in the context of the recently developed embedding approach for point defects [21] in which one describes the immediate neighborhood of the defect at a more accurate quantum mechanical level than the farther neighborhood. However, Ma *et al.*'s [21] goal is to describe the defect many-electron levels, which cannot necessarily be described with a single determinant many-body wave function, while ours is more simply to obtain the one-electron levels in a more accurate than DFT approach.

II. COMPUTATIONAL APPROACH

A. *GW* background

In many-body-perturbation theory, quasiparticle excitation energies are given by the equation

$$\left[-\frac{1}{2}\nabla^2 + v_N(\mathbf{r}) + v_H(\mathbf{r}) \right] \Psi_i(\mathbf{r}) + \int d^3r' \Sigma_{xc}(\mathbf{r}, \mathbf{r}', E_i) \Psi_i(\mathbf{r}') = E_i \Psi_i(\mathbf{r}), \quad (1)$$

where we use Hartree atomic units ($\hbar = e = m_e = 1$), v_N is the nuclear potential, and v_H the Hartree potential, Σ_{xc} the exchange-correlation self-energy, and Ψ_i is the quasiparticle wave function. In the *GW* approximation, the latter is calculated from $\Sigma(12) = iG(12)W(1^+2)$, where 1 is a shorthand for $\{\mathbf{r}_1, \sigma_1, t_1\}$, i.e., position, spin, and time of the particle 1, 1^+ means $\lim_{\delta \rightarrow +0} t_1 + \delta$, $G(12)$ is the one-electron Green's function, and $W(12)$ the screened Coulomb interaction. The exact one-particle Green's function is defined by $G(12) = -i\langle N|T[\psi(1), \psi^\dagger(2)]|N\rangle$, with T the time-ordering operator, $\psi(1)$ the annihilation field operator, and $|N\rangle$ the *N*-electron ground state. The screened Coulomb interaction is given by $W(12) = v(12) + \int d(34)v(3)P(34)W(42)$, and $P(12) = -iG(12)G(21)$ is the irreducible polarization propagator. In practice, it is usually obtained starting from some effective independent-particle Hamiltonian,

$$H^0 = -\frac{1}{2}\nabla^2 + v_N(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}), \quad (2)$$

with, for example, the LDA exchange-correlation potential $v_{xc}(\mathbf{r})$. The Green's function G^0 is then constructed from the eigenvalues ϵ_i and eigenfunctions ψ_i of

$$H^0 \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (3)$$

as follows:

$$G^0(\mathbf{r}, \mathbf{r}', \omega) = \sum_i \frac{\psi_i(\mathbf{r}) \psi_i^*(\mathbf{r}')}{\omega - \epsilon_i + i\delta \text{sgn}(\epsilon_i - \mu)}, \quad (4)$$

with μ the chemical potential and ω the energy variable of the Green's function. These LDA eigenstates form a convenient basis set in which the Green's function $G_{ij}^0(\omega) = \delta_{ij}(\omega - \epsilon_i \pm i\delta)^{-1}$ is diagonal. For a solid, the states are labeled by $i = \{n, \mathbf{k}\}$ with n a band index and \mathbf{k} the point in the Brillouin zone. The screened (W^0) and bare (v) Coulomb interactions and the polarization P^0 , on the other hand, are expressed in an auxiliary basis set of Bloch functions. In the LMTO implementation of the GW method, these are constructed from products of angular momentum partial waves [22] inside the spheres and plane waves confined to the interstitial space, which is afterward reduced to avoid linear dependence and rotated so as to diagonalize the bare Coulomb interaction [9,23,24]. We label them $E_{q\mu}(\mathbf{r})$ and the matrix of the Coulomb interactions in terms of them is then written

$$v_{\mu\nu}(\mathbf{q}) = \int d^3r d^3r' E_{q\mu}^*(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} E_{q\nu}(\mathbf{r}') \quad (5)$$

and

$$W_{\mu\nu}^0(\mathbf{q}, \omega) = [1 - v_{\mu\lambda}(\mathbf{q}) P_{\lambda\kappa}^0(\mathbf{q}, \omega)]^{-1} v_{\kappa\nu}(\mathbf{q}) \quad (6)$$

is obtained from a matrix inversion once $P_{\mu\nu}^0(\mathbf{q}, \omega)$ is known. The latter is also calculated directly in terms of the $\psi_{n\mathbf{k}}$ and eigenvalues $\epsilon_{n\mathbf{k}}$ [9]. In the above equation, summation convention over repeated indices is understood. In most other works, plane waves are used instead as basis functions. Let's write the rotation from the auxiliary functions E_{μ}^q to the LDA eigenstates as $\langle \psi_{k_n} | \psi_{\mathbf{k}-\mathbf{q}n'} E_{\mu}^q \rangle$. Using this rotation matrix, we can express W as

$$W_{mn'mk}^0(\mathbf{q}, \omega) = \sum_{\mu\nu} \langle \psi_{k_n} | \psi_{\mathbf{k}-\mathbf{q},n'} E_{\mu}^q \rangle W_{\mu\nu}^0(\mathbf{q}, \omega) \times \langle E_{\nu}^q | \psi_{\mathbf{k}-\mathbf{q}n'} | \psi_{k_m} \rangle. \quad (7)$$

The self-energy matrix is then given by

$$\Sigma_{nm}(\mathbf{k}, \omega) = \frac{i}{2\pi} \int d\omega' \sum_{\mathbf{q}} \sum_{n'} G_{nn'}^0(\mathbf{k} - \mathbf{q}, \omega - \omega') \times W_{nn'm;\mathbf{k}}^0(\mathbf{q}, \omega') e^{i\delta\omega'}, \quad (8)$$

in which we recognize the schematic $\Sigma = iGW$ but which makes it clear that to obtain the energy and \mathbf{k} -space dependent form, a triple convolution is involved over energy ω' , \mathbf{q} and band index n' .

This self-energy matrix is energy dependent and contains, in principle, information, not only on the quasiparticle energies but also the satellites and noncoherent parts of the one-electron excitation. In the QSGW method, we now reduce this to a nonlocal but energy-independent and Hermitian

matrix:

$$\tilde{\Sigma}_{nm}(\mathbf{k}) = \frac{1}{2} \text{Re}[\Sigma_{nm}(\mathbf{k}, \epsilon_{n\mathbf{k}}) + \Sigma_{nm}(\mathbf{k}, \epsilon_{m\mathbf{k}})]. \quad (9)$$

We now define $\Delta \tilde{\Sigma}_{nm}(\mathbf{k}) = \tilde{\Sigma}_{nm}(\mathbf{k}) - v_{nm}^{xc}(\mathbf{k})$, where we subtract the matrix element of the LDA exchange-correlation potential taken between the Bloch eigenstates. We can then add this correction to the exchange-correlation potential to the H^0 Hamiltonian and re-diagonalize the latter to find new independent particle eigenvalues and eigenstates and repeat the cycle of calculating $\tilde{\Sigma}$. At the convergence of this iteration, the eigenvalues of the final H^0 , which we will call H^{QSGW} are identical to the quasiparticle energies. We may view this as finding a G^0W^0 perturbation theory solution of Eq. (1) starting from H^0 but refining H^0 so the perturbation becomes negligible. In this sense, the quasiparticle energies are self-consistent and independent of the starting approximation but they are still real and do not provide information on the lifetime of the actual quasiparticle states.

B. Bloch-function and real space representation

The Bloch functions of the H^{QSGW} are now known as an expansion in the LMTO basis set,

$$|\psi_{n\mathbf{k}}\rangle = \sum_{\mathbf{R}i} |\chi_{\mathbf{R}i}^{\mathbf{k}}\rangle b_{n\mathbf{R}i}^{\mathbf{k}}, \quad (10)$$

so we can re-express the self-energy correction matrix as $\Delta \tilde{\Sigma}_{\mathbf{R}i, \mathbf{R}'i'}(\mathbf{k})$. Here \mathbf{R} label the sites in the unit cell and i the muffin-tin orbitals. The latter are labeled by angular momentum quantum numbers (l, m) as well as a third index, labeling the choice of smoothed Hankel function decay and/or local orbital (confined to the muffin-tin sphere). See Ref. [9] for a full description of the full potential (FP)-LMTO method used. Finally, performing an inverse Bloch sum or Fourier transform, we obtain $\Delta \tilde{\Sigma}_{\mathbf{R}, i; \mathbf{R}'+\mathbf{T}, i'}$ fully expressed in real space, where \mathbf{T} are the lattice translation vectors.

Next, let us consider the same self-energy of the bulk system but represented in a supercell. Obviously there is a one-to-one mapping $\{\mathbf{R}, \mathbf{T}\} \leftrightarrow \{\mathbf{R}_S, \mathbf{T}_S\}$ between positions \mathbf{R} in the primitive cell with lattice vectors \mathbf{T} to the positions inside the supercell, \mathbf{R}_S , and the superlattice's lattice vectors, where $\mathbf{T}_S = \sum_i^3 n_i \mathbf{A}_i$ with n_i integers, and the superlattice is defined by $\mathbf{A}_i = \sum_j N_{ij} \mathbf{a}_j$ with N_{ij} a set of integers. Thus, we can obtain $\Delta \tilde{\Sigma}_{\mathbf{R}_S, i; \mathbf{R}'_S + \mathbf{T}_S, i'}$ by a simple relabeling procedure. In practice, these are stored for every \mathbf{R}_S using a neighbor table out to some maximum distance $|\mathbf{R}'_S + \mathbf{T}_S - \mathbf{R}_S| \leq d_{\text{max}}$.

C. Self-energy cut-and-paste approach

The overall scheme for constructing the defect cell and its self-energy is as follows. For example, let us consider a 64-atom supercell to model the defect. We then start by creating the self-energy matrix of the perfect supercell (the host) from the self-energy in the primitive cell in the above form in real space and labeled according to the 64-atom cell $\{\mathbf{R}_S, \mathbf{T}_S\}$ scheme. Once we have the self-energy matrix in the host supercell, we need to replace it by that for the defect within a certain range d_{def} from the defect atom. For that purpose, we construct a smaller supercell containing the defect, which we ultimately plan to use, e.g., an eight-atom cell and carry

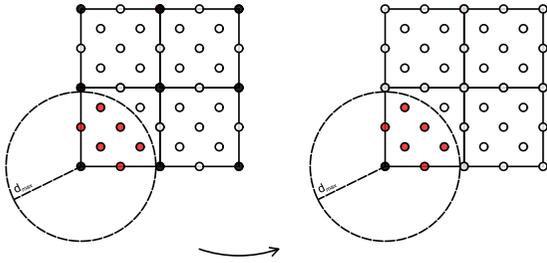


FIG. 1. Schematic illustration of the self-energy editor cut-and-paste method. The left shows the 64^{8d} supercell with defect atoms shown as black spheres and host atoms as open circles and the atoms within a range d_{\max} from the defect atom indicated by the dashed circle are red. The target system 64^{1d} with 1 defect is shown on the right. In this example, we assume only the defect atom itself comprises the defect region and the atoms within the range d_{\max} contribute to the self-energy in real space. The self-energy $\tilde{\Sigma}_{\mathbf{R},\mathbf{R}'}$ of the atom pairs corresponding to the red atoms connected to the defect atom in the target 64^{1d} cell are replaced by those from the 64^{8d} supercell, shown on the left.

out a self-consistent DFT calculation for it and subsequently a QSGW calculation of its self-energy. We then transform the self-energy of this defect cell again to a new 64-atom cell by a similar relabeling step. Let us call this the 64^{8d} cell where 64 indicates the number of atoms in the cell and the superscript indicates the cell contains eight defects. Next, we modify the host 64-atom cell by inserting the defect. We then carry out a self-consistent calculation for it at the DFT level and construct its self-energy by the following cut-and-paste method. The first step of the method is to create a blank place holder in memory to hold the self-energy around each atom, the type of atom, and their orbitals at each of its neighbors. We then copy the corresponding self-energy into it from the host 64-atom cell and subsequently replace it by that of the defect 64^{8d} cell for atoms within a certain distance d_{def} from the defect site. The rest of the atoms are left unchanged as host atoms. Each copy step only happens according to the neighbor table up to a maximum range d_{\max} . The copying of the self-energy orbital blocks happens by pairs. Only half of the $\Sigma_{\mathbf{R}_S i; \mathbf{R}'_S + \mathbf{T}_S, i}$ matrix elements need to be constructed because of the Hermiticity. Once assembled in real space, it can be Fourier transformed back according to the periodicity of the supercell to find $\Sigma_{\mathbf{R}_S i; \mathbf{R}'_S i}(\mathbf{k}_S)$. Finally, we then need to carry out just one DFT self-consistent step in which the thus assembled estimate of the self-energy is added to the H^0 DFT Hamiltonian and we can evaluate its band structure. The scheme is illustrated in Fig. 1.

It is clear that for the scheme to work, d_{\max} must fit within the small defect-containing supercell, so the self-energy in the final cell for a pair of atoms of which one is within the defect range d_{def} does not have a neighbor of the wrong type, in other words, it must still be a hostlike atom as in the final cell, not another defect atom.

Furthermore, we need to allow for relaxation of the atoms near the defect. Therefore, the atoms are mapped between the different cells based on their atom numbering and connectivity, not on the basis of their exact position. In principle, one could first relax the atoms in the large cell with the defect

at the DFT level and then do the mapping to the perfect crystal and small defect cell even if their atomic positions do not perfectly match. Alternatively, we could assume that the self-energy is not too sensitive to the relaxation and hence keep the self-energy fixed after our initial cut-and-paste operation and, afterward, relax (or relax again) the positions in the presence of that fixed self-energy. It is important here to remember that the self-energy provides only a correction to the electronic structure beyond the DFT Hamiltonian. The main defect-induced changes are already contained at the DFT level.

D. Computational details

The method has been implemented in the LMTO and QSGW suite of codes, named Questaal (Quasiparticle Electronic Structure and Augmented LMTOs) [25,26]. The basis sets are specified in terms of the angular momentum cutoffs and smoothed Hankel functions. For the initial tests on GaAs with the As_{Ga} antisite defect, we used a rather minimal basis set as specified, along with the results. This leads to an overestimate of the band gap but is convenient for our present purpose of demonstrating the validity of the approach. Lattice positions are optimized for the crystal cells with defects. Details of the supercells chosen in the cut-and-paste approach are given along with the test results. In the final Sec. III D, we use a larger basis set to achieve accurate comparison to experiment.

III. RESULTS

A. Gap convergence with self-energy real-space cutoff

We start by testing the core idea of a finite-range self-energy for bulk GaAs. Table I shows the band gap of GaAs in QSGW as function of the cutoff d_{\max} used in the real space representation of the self-energy. We can see that as soon as $d_{\max}/a > 1$ with a the lattice constant or a cluster of about 30 atoms is included, the gaps become reasonable, although the convergence is not uniform and it takes until a cluster of about

TABLE I. Convergence of band gap with d_{\max} in GaAs with Basis set Ga: *spd*, As: *spd*. $a = 10.66$ Bohr is the lattice constant in the zinc blende structure. The gap of GaAs the LDA gap with this basis set is 0.50 eV and the k-space QSGW gap is 2.35 eV.

d_{\max}/a	No. of neighbors	Gap (eV)
0.6	5	1.84
0.8	17	1.76
1.0	29	2.09
1.2	47	2.26
1.4	87	2.30
1.6	147	2.29
1.8	191	2.30
2.0	275	2.32
2.2	345	2.32
2.4	417	2.33
2.6	457	2.34
3.0	461	2.35
∞		2.35

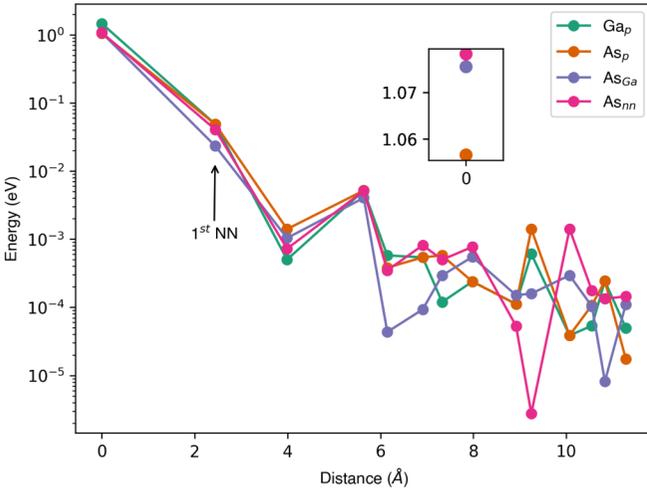


FIG. 2. The trace of submatrices of corresponding atom-atom pair $\text{Tr}[\Sigma_{\mathbf{R}_S i; \mathbf{R}'_S + \mathbf{T}_S i'}]$ in the self-energy matrix is plotted as function of distance $|\mathbf{R}'_S + \mathbf{T}_S - \mathbf{R}_S|$. The basis atoms of primitive GaAs crystal (Ga_p and As_p), the defect (AS_{Ga}) from the 64-atom cell and its nearest neighbor (AS_{nn}) are chosen to illustrate the drastic drop in energy corrections with distance. Both structures are in $q = 0$ state. In the inset on the top, the on-site energy corrections of the As atoms are seen more clearly.

450 atoms or $d_{\text{max}}/a \approx 3$ to get absolute convergence. We note that in an eight-atom supercell, the distance between defects is only one cubic lattice constant a and thus choosing $d_{\text{max}} = a$ would still result in some $\Sigma_{\mathbf{R}, \mathbf{R}' + \mathbf{T}}$ with both \mathbf{R} and $\mathbf{R}' + \mathbf{T}$ sites being defect sites, which is incorrect for the dilute limit. One might think one needs to restrict $d_{\text{max}} \leq a/2$ to avoid such defect-defect matrix elements of the Σ matrix. However, such a small choice of d_{max} would lead to an unacceptable error in the band gap. We will show, however, that using $d_{\text{max}} = a$ even in an eight-atom cell introduces a negligible error.

The short-distance nature of the self-energy in real space is illustrated in Fig 2. We here show the trace of each block of the self-energy matrix for pairs $\{\mathbf{R}, \mathbf{R}' + \mathbf{T}\}$ as function of their separation distance. While the self-energy operator $\tilde{\Sigma}(\mathbf{r}, \mathbf{r}')$ itself, in principle, falls off as $1/|\mathbf{r} - \mathbf{r}'|$, being a screened exchange type term, the matrix elements $\Delta \tilde{\Sigma}_{\mathbf{R}_i, \mathbf{R}' + \mathbf{T}_i'} = \langle \chi_{\mathbf{R}_i} | \tilde{\Sigma}(\mathbf{r}, \mathbf{r}') - v_{\text{xc}} | \chi_{\mathbf{R}' + \mathbf{T}_i'} \rangle$ fall off much faster because they are dominated by the overlap of the corresponding basis function orbitals. The on-site elements are clearly seen to be two to three orders of magnitude larger than the intersite elements with second or higher nearest neighbors. They oscillate somewhat as we go to further neighbors. Specifically, we can see, that if we cut off these matrix elements at a distance $d_{\text{max}} > a \approx 5.6 \text{ \AA}$, the off-site matrix elements of the third neighbor (at distance a) is about 100 times smaller than the on-site element and 10 times smaller than the first nearest-neighbor one. Furthermore, in the perfect crystal, this is a Ga-Ga or As-As type matrix element shown, respectively, by the green or orange symbols, while in the 64-atom defect cell, the matrix elements connected to the AS_{Ga} defect site (shown in blue) at the origin would be a As-Ga matrix element but it is not significantly different from the perfect crystal one. The error we make by substituting this

from the eight-atom cell is that we use an $\text{AS}_{\text{Ga}}\text{-AS}_{\text{Ga}}$ third-neighbor matrix element instead of $\text{AS}_{\text{Ga}}\text{-Ga}$ matrix element of the 64-atom cell, but these differ by only $8 \times 10^{-4} \text{ eV}$. This shows that even though the eight-atom cell may seem too small to respect the $d_{\text{max}} = a$ cutoff, it will introduce only a negligible error even when we have a wrong type of atom matrix element. We might even have gotten away with a larger d_{max} using an eight-atom cell. While this would introduce more wrong types of atom matrix elements, and these start deviating more from each other relatively, they are all of order 10^{-3} eV and would not contribute much to the final result.

From the inset in Fig. 2, we can see that even the on-site self-energy of As in the perfect crystal or in the defect site are very close to each other. The essential point in correcting the self-energy matrix of the perfect crystal represented in the supercell, which already incorporates the host band-gap change, is to replace that of the Ga atom by an As atom. That on-site matrix element (green circle versus pink circle) makes a difference of order of a few 0.1 eV and is sizable. The intersite elements of the self-energy beyond nearest neighbors are so much smaller that they play only a minor role and this explains why we can restrict the range of the self-energy matrix elements rather severely with a small d_{max} . This also provides *a posteriori* an explanation for why schemes such as LDA + U [11,12] for modifying the band structure or other local on-site corrections [10,14] have had considerable success in adjusting the gap in defect calculations.

We finally note that the self-energy matrix element localization and the smoothness of the decay could be further improved by means of more localized screened muffin-tin-orbitals such as the jigsaw orbitals proposed in Ref. [25]. It was already shown there that the band gap as a function of self-energy cutoff decays quickly and smoothly using these more localized basis functions than with standard muffin-tin orbitals.

B. Basic properties of the AS_{Ga} antisite defect

Next, we check the viability of the scheme for the case of the AS_{Ga} antisite in GaAs. This is a well-studied defect, known as the EL2 defect, or at least closely related to it [27,28]. In the $q = 0$ state, it has a single a_1 defect level filled in the gap and a single t_2 empty resonance just above the CBM. The excited state $a_1^1 t_2^1$ is twofold degenerate and both its $S = 1$ and $S = 0$ configurations are orthogonal to the ground state $a_1^2 t_2^0$. This degeneracy leads to a symmetry breaking distortion, such that the antisite As atom is pushed through the interstitial position and the initial single point defect turns into Ga-vacancy and As-interstitial $V_{\text{Ga}} + \text{As}_i$ defect complex. Due to this displacement, the system might get trapped in a metastable state at lower energy than the excited state but at the distorted geometry. This metastable state is labeled $1a^0 2a^2$ where now the levels are labeled according to the C_{3v} distorted geometry. The ground state in this notation is $1a^2 2a^0$ and the excited state is $1a^1 2a^1$. In other words, $1a$ in C_{3v} derives from the a_1 in T_d while $2a$ derives from t_2 [28]. This state is associated with the photoquenching behavior of EL2 defect in Ref. [28]. Therefore, the optical excitation energy from the a_1 to the t_2 level is of interest and can be directly related to the corresponding QSGW levels since it occurs from

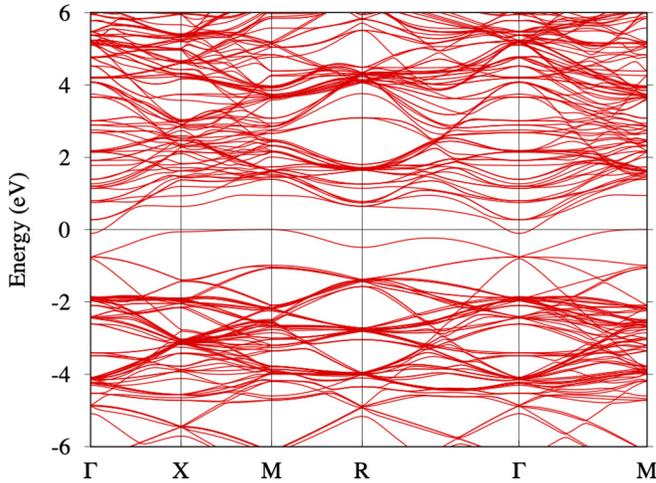


FIG. 3. Band structure of As_{Ga} in $q = 0$ state in GaAs in 64-atom cell in LDA. This supercell has simple cubic form and the high-symmetry \mathbf{k} -points are $\Gamma = (0, 0, 0)$, $X = (1, 0, 0)$, $M = (1, 1, 0)$, and $R = (1, 1, 1)$ in units $2\pi/a_s$ with $a_s = 2a$ the supercell lattice constant.

the ground state $q = 0$ geometry. Furthermore, the Green's function calculations of Bachelet *et al.* [27] provide detailed information on the position of the Kohn-Sham defect levels of a_1 and t_2 symmetry in two charge states. We identify these two defect levels in defect supercell band structure and compare our results with Refs. [27,28]. The ground-state defect level of neutral EL2 is well known experimentally and we discuss this further in Sec. III D.

C. Application of the method to As_{Ga}

In Fig. 3, we show the band structure of a 64-atom supercell containing an As_{Ga} antisite defect at its origin in the LDA and for $q = 0$ state. In this paper, we used a minimal basis set of only a single κ and smoothing radius R_{sm} , spd orbital set on Ga and As. The gap is thereby underestimated as 0.65 eV in LDA and overestimated as 2.25 eV in QSGW. This should facilitate recognizing the defect level. Nonetheless, we see in Fig. 3 that the electronic structure of the defect and even the gap are barely recognizable. The defect band is the highest occupied band but is seen to be so much broadened that, in combination with the LDA gap underestimate, it touches the valence band maximum (VBM) at the Γ point. The CBM occurs 0.65 eV above it at the Γ point in LDA, while a converged basis set would give an even lower LDA gap of only 0.51 eV. Clearly, the band structure of this system cannot be examined accurately with DFT level calculations.

Next, we construct the defect in an eight-atom cell and perform a QSGW calculation for it. From the corresponding band structure shown in Fig. 4, it is clear that this cell is much too small to adequately represent the defect electronic structure. In this figure, it is not even clear which is the defect band and which is the lowest conduction band. On the other hand, we will show that this cell is sufficient to reconstruct the self-energy components in the immediate neighborhood of the defect.

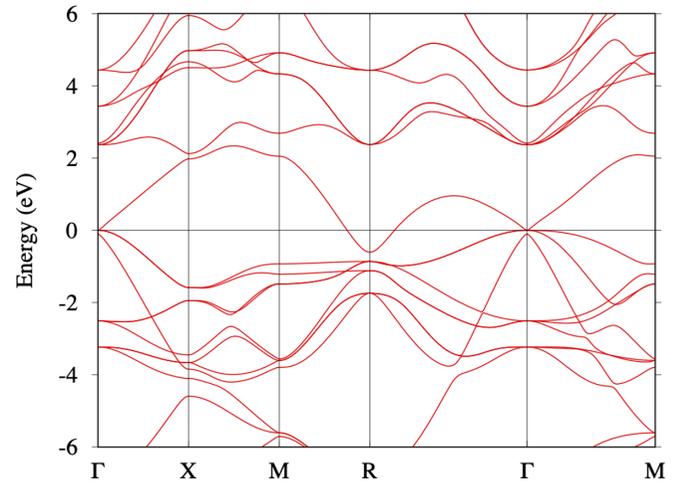


FIG. 4. Band structure of As_{Ga} in eight-atom GaAs cell in QSGW. This supercell is also a simple cubic but with lattice constant a . The high-symmetry points are labeled the same as in Fig. 3 but correspond to a Brillouin zone twice as large in each direction.

In Fig. 5, we show the band structure of the 64-atom cell obtained by means of our self-energy cut-and-paste approach with the dashed red lines. The defect region contains only the defect atom itself and the self-energy range d_{max} was set to one lattice constant (5.64 Å). In the same figure, we show the fully self-consistent QSGW results in the same 64 atom cell with the solid black lines after aligning the two at the VBM. The dispersion of this defect band is about 0.7 eV and results from periodically repeated defects in the 64-atom supercells, so from defects that are $2a_{\text{cell}}$ or 11.28 Å apart. Comparing with the band structure in a larger cell in Fig. 8, we can see that the defect bandwidth is mostly reduced at R and Γ but the top of the band near X and M stays the same. We therefore identify the eigenvalue near its maximum with the isolated

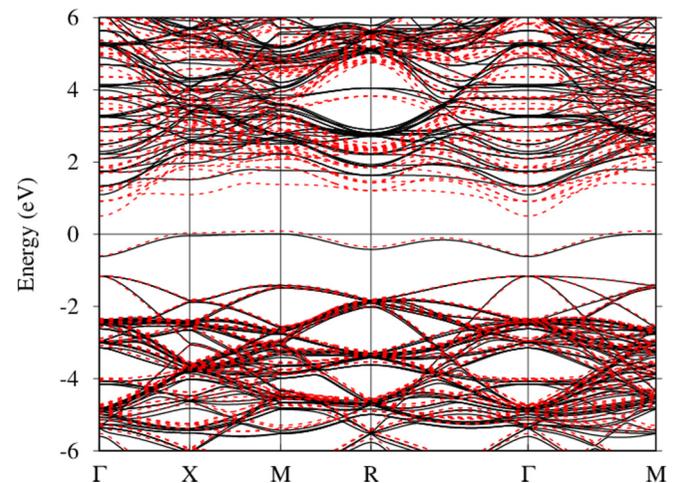


FIG. 5. Band structure of As_{Ga} in $q = 0$ state in GaAs in 64-atom cell: Fully QSGW (solid black lines), with self-energy constructed by cut-and-paste approach (red dashed lines) using only the defect atom as defect region aligned at the VBM. The zero of energy is the Fermi level for the defect system in the full QSGW case for the neutral defect.

defect level. The fact that the defect band dispersion does not follow the expected form $E_d + 2t[\cos(k_x a) + \cos(k_y a) + \cos(k_z a)]$ for a simple isotropic s -band in a nearest-neighbor simple cubic lattice, where the maximum would be at R and the minimum at Γ , indicates that the interactions between defect states are not isotropic because of the underlying crystal structure. The filled defect band is now clearly detached from the VBM and it occurs at about 1.12 eV above the VBM. Its position above the VBM and even its dispersion is in excellent agreement between the cut-and-paste and fully self-consistent results. Furthermore, this defect-level position above the VBM is in good agreement with Bachelet's Green's function calculation [27] of an isolated defect, which gives 1.23 eV for the $q = 0$ state.

The CBM is also clearly seen (overestimated as 2.25 eV above the VBM by our small basis set) in the full self-consistent calculation. This gap is somewhat smaller in the cut-and-paste case (1.66 eV). This is because the range cutoff d_{\max} applied to Σ reduces the gap. Second, in this 64-atom cell, we do not yet approach the dilute limit where the perfect crystal gap with this cutoff (2.09 eV) would be recovered. This can be attributed to the t_2 defect level interacting with the CBM. In both cases, we can identify the second defect level, the t_2 state, which here is a resonance and can be recognized as a flat band near X at about 1.55 eV above the Fermi level, which coincides with the top of the a_1 defect band, in direct QSGW calculation and 1.07 eV in the cut-and-paste method. This defect level lies 0.42 eV above the CBM in direct QSGW calculation and 0.60 eV in the cut-and-paste approach. The energy splitting between the t_2 and a_1 state taken as the energy difference at the X point, because of above reasons, hence 1.55 eV in direct calculation and 1.06 eV in our approach. Dabrowski and Scheffler [28] reports this splitting as 1.18 eV experimentally and as 0.97 eV at the DFT-level calculations, while in the Green's function calculation it is 0.87 eV [27]. Our result for this splitting of the $t_2 - a_1$ level is thus comparable in accuracy with the previous calculations and in fact closer to the experimental value, which, as mentioned earlier, is important for understanding the optical behavior of this defect. Hence, the cut-and-paste method is found to be a viable approach.

Next, we test whether replacing only the self-energy related to the defect atom itself is sufficient or whether we need to include a larger defect cluster region. In our cut-and-paste method, including the nearest neighbors in the defect region corresponds to taking the nearest neighbor of the defect atom as another center for the $\Delta\tilde{\Sigma}_{\mathbf{R},\mathbf{R}'+\mathbf{T}_s}$ to be taken from the small defect-containing (eight-atom) cell. This approach could be thought to be problematic because it extends the range of this $\Delta\tilde{\Sigma}$ with \mathbf{R} being a nearest neighbor of the defect beyond its unit cell, and hence includes pairs connecting this atom to other defect atoms, while in the dilute limit or in the large supercell with a single defect, there should only be one defect atom. However, for our eight-atom cell and with $d_{\max} = a$ this problem already occurred even for the defect atom itself but, as already discussed in Sec. III A, the off-site matrix elements fall off very quickly and this is found not to be a serious problem. With this approach, we observe that the valence and defect bands remain in the same position, but the CBM value drop by 0.14 eV, which deviates more from the QSGW

TABLE II. The energy band gap, E_g , the defect level at the X -point with respect to VBM, $a_{1\text{VBM}}^X$, the defect-level splitting at the X point, $(t_2 - a_1)^X$, and the a_1 defect band width $w(a_1)$, are presented for various schemes in $q = 0$ (top half) and in $q = 2$ (bottom half) states. The subscript numbers represent the small cell atom number which was used to build up the final cell in the cut-and-paste method. The defect levels for the $q = +2$ charge state include a background-image charge correction as discussed in the text of -0.29 eV for 64-atom and -0.19 eV for the 216-atom cell.

	E_g	$a_{1\text{VBM}}^X$	$(t_2 - a_1)^X$	$w(a_1)$
Dabrowski and Scheffler [28]		0.6	0.97	
Bachelet <i>et al.</i> [27] (LDA)	0.7	1.23	0.87	
64-atom QSGW	2.25	1.12	1.55	0.62
64 ₈ ^a	1.66	1.19	1.07	0.68
64 ₈ ^b	1.52	1.19	1.04	0.69
64 ₃₂ ^a	1.69	1.01	1.02	0.67
64 ₃₂ ^b	1.64	0.98	1.03	0.68
216 ₈ ^a	1.72	1.14	1.24	0.19
216 ₈ ^b	1.67	1.14	1.23	0.19
Bachelet <i>et al.</i>		0.69	0.99	
64-atom QSGW	2.24	0.40	1.60	0.60
64 ₈ ^a	1.71	0.77	0.95	0.67
64 ₈ ^b	1.58	0.78	0.94	0.67
64 ₃₂ ^a	1.63	0.63	0.96	0.67
64 ₃₂ ^b	1.64	0.63	0.95	0.67
216 ₈ ^a	1.70	0.42	1.24	0.16
216 ₈ ^b	1.65	0.43	1.24	0.16

^aDefect region is defect atom itself.

^bDefect region is defect atom and its nearest neighbors.

calculation than with only the defect atom itself. These results are shown in the Supplemental Material [29].

To describe the neighborhood of the defect atom more accurately and work with a larger cutoff in a safe way, it would be advantageous to enlarge the size of the cell from which the defect and its neighbors' self-energy is extracted. This might then also allow us to increase the d_{\max} or range of the self-energy cutoff. To further test the convergence of our scheme, we now consider a somewhat larger cell than the eight-atom cell to extract the defect atom and its neighbors' self-energy. For this purpose, we incorporate the defect in a 32-atom supercell of GaAs. The d_{\max} value is set to the lattice constant of this cell, $1.73a$ of the conventional cell, which is then also the nearest defect distance.

Building the host cell using this small cell does not provide considerable improvement for the cut-and-paste method. This is expected since we already know that the self-energy matrix elements fall off rapidly with intersite distance. Furthermore, the exact QSGW calculation of this cell is now a more expensive calculation and not so much is gained by increasing the size to a 64-atom cell using the cut-and-paste approach. Nonetheless, we test its performance to check the convergence of our cut-and-paste approach in terms of the size of the defect region and self-energy cut-off distance d_{\max} . The results of this scheme are summarized in Table II with the label 64₃₂. Different defect region descriptions and charge states were also considered. In Fig. 6, we compare the two cut-and-paste band

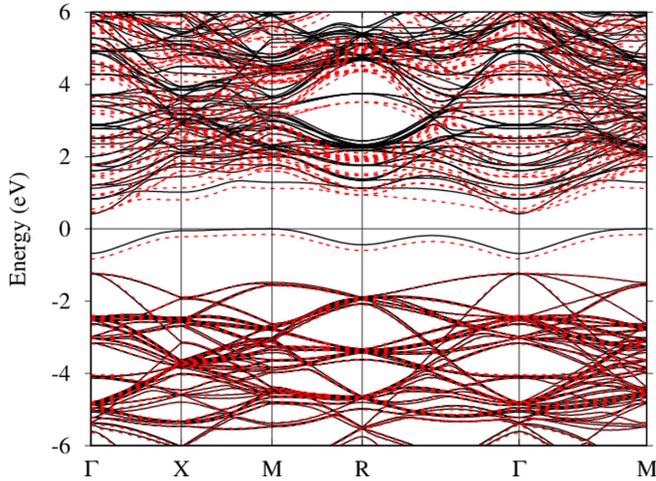


FIG. 6. Band structure of As_{Ga} in $q = 0$ state in GaAs in 64-atom cell: 64_8 scheme (solid black lines), and 64_{32} scheme (red dashed lines), both with self-energy constructed by cut-and-paste approach using only the defect atom as defect region, aligned at the VBM. The zero of energy is the Fermi level for the defect system in 64_8 scheme.

structures of the defect containing a 64-atom cell in the $q = 0$ state, built from the eight-atom cell (64_8) and from the 32-atom cell (64_{32}). The almost identical bands show that there is nothing to be gained in the accuracy of the method by extracting the defect atom self-energy from a 32-atom compared to an eight-atom cell and is furthermore significantly more computationally expensive. We conclude that the most effective approach for the cut-and-paste method is building the defect-containing host cell from the eight-atom cell and it is sufficient to define the defect region as the defect atom itself. The results obtained by this approach are in good agreement with our exact QSGW calculations and previous studies in the literature.

Next we want to address whether the cut-and-paste approach also works for charged defect states and correctly describes the trend with charge observed in the Green's function calculations. In Fig. 7, we show the band structure of the same cell, but with the defect in the $q = 2$ state. In this case, the exact QSGW band gap does not change, but the defect level moves closer to the VBM. The cut-and-paste approach yields 0.05 eV larger band gap. Although the defect band dispersion is again reproduced faithfully, the position of defect level deviates by about 0.2 eV between the cut-and-paste and exact calculation. This is still an acceptable precision of the cut-and-paste method and indicates that the nearly perfect agreement seen earlier for the $q = 0$ state is perhaps somewhat coincidental.

On the other hand, the difference between the two charge states regarding the defect level position agrees with prior work and has a clear physical meaning. It can be related to the different final geometries after atomic relaxations in different charge states. For the $q = 0$ state, we observe outward breathing of nearest As atoms, such that they move further from the defect atom about 4% of the nearest-neighbor distance. Dabrowski and Scheffler [28] reported similar lattice relaxation effects. On the other hand, in the $q = 2$ state, the nearest-neighbor distance remains the same as in the perfect crystal.

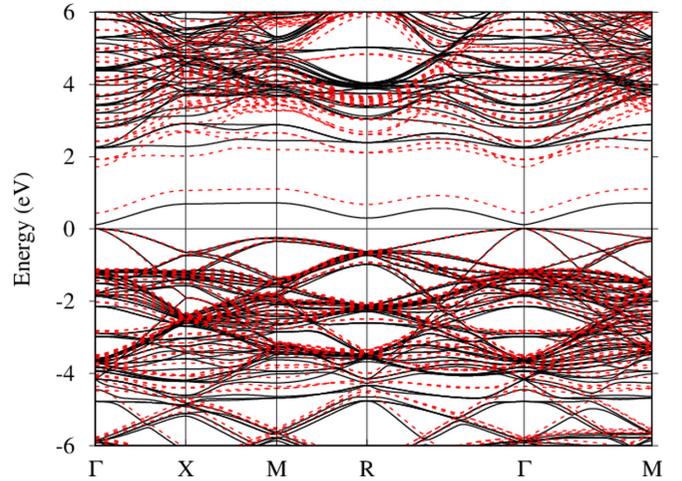


FIG. 7. Band structure of As_{Ga} in $q = 2$ state in GaAs in 64-atom cell: fully QSGW (solid black lines), with self-energy constructed by cut-and-paste approach (red dashed lines) using only the defect atom as defect region, aligned at the VBM. The zero of energy is the Fermi level for the VBM in the full QSGW case for the doubly ionized defect.

There is an additional issue with charged defect levels, namely, the spurious interaction of the image charges with the compensating homogeneous background charge density. Such corrections are well known to be required for defect formation energies or total energies and can be estimated in a point charge model as $E_{\text{cor}} = -\alpha q^2 / 2L\epsilon$, where α is a Madelung constant for the supercell, L the linear size of the supercell and ϵ the dielectric constant of the medium. Closely related, there is a correction in the potential for a one-electron level which is $\delta V = -\alpha q / L\epsilon$. Note that this correction is linear in the charge state but does not include the factor of 2 in the denominator. For $q = 2$, the two corrections are the same. Using the dielectric constant of GaAs of 12.9 and the size of our supercell, this amounts to a correction of -0.29 eV for the 64-atom cell and -0.19 eV for the 216-atom cell. This gives a difference between the $q = 0$ and $q = +2$ defect levels of 0.43 eV for the 64-atom cell using cut-and-paste.

One important advantage of our method is that once the exact QSGW calculation for the small cell is done, moving to larger host cells and calculating the defect properties in the dilute limit is a straightforward procedure and one can achieve *GW*-level accuracies for the cost of an LDA calculation for even larger cells. Although we do not include the exact calculations for these larger cells in this paper, we analyze a 216-atom cell, i.e., a $3 \times 3 \times 3$ supercell of the conventional simple cubic eight-atom cell. In Fig. 8, we see that in the 216-atom cell dispersion of the defect band is almost zero, indicating we are close to the dilute limit. Furthermore, this allows us to better evaluate the nature of the defect band dispersion. As mentioned earlier, the top of the defect band which is flat between X and M remains the same as in the smaller 64- and 32-atom cells but the bandwidth is reduced at Γ . This helps us to identify the top of the defect band with the dilute limit defect level.

It is important to point that obtaining this 216-atom cell band structure only required an additional LDA calculation

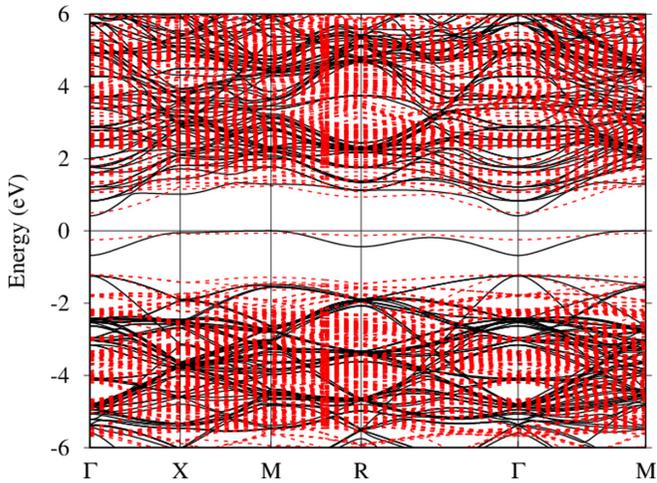


FIG. 8. Band structure of As_{Ga} in $q = 0$ state: 64_8 scheme (solid black lines), and 216_8 scheme (red dashed lines), aligned at the VBM. Defect region is defined only with the defect atom. Dispersion of a_1 level vanishes as defect-defect distance increases in a fashion explained with the tight-binding model above. Note that the Brillouin zones of the 216- and 64-atom cells have the same symmetry labeling but the latter is larger and contains fewer bands. The $\Gamma - X$ distance of the 216 atom cell is scaled down by a factor 2/3 compared with the 64-atom one and similarly for the other directions.

of this cell. Once the eight-atom cell GW calculation is done, moving to even larger cells only requires the LDA calculation of the desired single defect containing supercell. This provides an immense reduction in computational cost. For instance, we can compare the computation time of exact QSGW calculations for 8-atom, 32-atom, and 64-atom cells. The convergence parameters for these calculations are kept the same and the \mathbf{k} -point meshes are kept equivalent by taking a smaller set with approximately the same grid-density in the larger cells. The number of processors for each calculation is adjusted proportional to the \mathbf{k} -point mesh grid. Under these circumstances, the fully consistent QSGW calculation was completed in 0.6 h for the 8-atom cell, 26.4 h for the 32-atom cell, and 181 h for the 64-atom cell. In other words, the computing timescales like N^η with $\eta \approx 1.8$ and N the number of atoms or 3η for the scaling with the linear size of the supercell. On the other hand, the LDA calculation of the 64-atom cell took less than half an hour. For this particular 64_8 scheme, the bottleneck of cut-and-paste method was the eight-atom cell QSGW calculation. For larger single defect cells, the LDA calculation might become the bottleneck due to large number of atoms, however, the computation cost would still be insignificant compared to the QSGW calculation of the same cell. For instance, the LDA calculation of the host cell in our 216_8 scheme was performed using resources comparable to that of 64_8 LDA step, and this step took around 6 h. Note that all the atoms are fully relaxed in this step, and one could reduce the computation time even further by only letting the first few nearest-neighboring atoms relax.

D. Large basis sets and comparison to experiment

To test the fidelity of the theory, we finally make a careful comparison against the experimentally observed neutral EL2

deep donor level ($E_v + 0.75$ eV) [28,30,31]. This defect level was first identified with the As_{Ga} antisite by Weber *et al.* [30] based on electron paramagnetic resonance (EPR) and activation of the unpaired spin +1 charge state from the neutral state by an optical transition to the conduction band, which was found to occur at 0.75 eV. This was later also confirmed by thermionic emission, capacitance spectroscopies and optical studies [31–33]. In view of the experimental gap of 1.52 eV at low temperature, this places the defect level almost exactly at mid gap. We revisit the neutral EL2 level with a reasonably well converged basis, and also include spin-orbit coupling (SOC). We repeat the procedure described above, using a 32-atom supercell, with an *spdfspd* basis on both the Ga and As, and local orbitals to include the Ga 3d in the valence. The basis also includes *sp* “floating orbitals” (smoothed Hankel functions without augmentation spheres [25]) centered at the two high-symmetry interstitial sites along the [111] line. The QSGW bandgap (1.7 eV, including SOC) is slightly less than that of a fully converged basis (1.8 eV). This includes a reduction in the bandgap of 0.11 eV from SOC. The gap reduction is expected, as the split-off valence band at Γ is 0.33 eV below the VBM, both experimentally and in QSGW.

We note that even with his well-converged basis set the bandgap is overestimated because QSGW under-screens W , because the RPA polarizability omits electron-hole attractions that connect the electron and hole parts of the bubble. If such attractions are included, e.g., via ladder diagrams, the bandgap is reduced to a value very close to the observed gap. It was discovered soon after QSGW was first formulated that, empirically, the RPA dielectric constant ϵ_∞ is uniformly 20% too small for a wide range of insulators, strongly correlated or not [34]; see in particular Fig. 1 in Ref. [35]. Adding ladder diagrams greatly improves on $\epsilon(\omega)$, as has been known from pioneering work in the groups of Louie [36] and Reining [37]. It has recently shown that improving W with ladders, and using this W in the QSGW self-consistency cycle, almost completely eliminates the overestimate of band gaps in weakly correlated semiconductors, and it also corrects for the underestimate of ϵ_∞ [38]. As an alternative, a hybrid approach has long been used, mixing LDA and QSGW [34]. Kotani and his coworkers showed that a hybrid of 80% QSGW and 20% LDA yields uniformly good band gaps in many weakly correlated semiconductors [39]. The hybrid approach is an inexpensive, albeit *ad hoc* way to mimic the effect of ladder diagrams, and we use it here to refine our estimate of the neutral EL2 level.

Figure 9 shows the density of states (DOS) of the 128 supercell within QSGW for two scenarios: the black data is the DOS for an ideal (unrelaxed) structure, while the red data shows the effect of relaxing the four nearest neighbors around the As_{Ga} only. Limiting relaxations to nearest neighbors simplifies the embedding procedure, and tests showed that more complete relaxations made minor further changes. It is seen that the lattice relaxation induces a shift in the defect level, moving it about 0.3 eV closer to the valence band.

To estimate the EL2 energy in the limit of an infinite cell, the center of gravity of the band was calculated. It is found to be close to the Fermi level at the top of the defect band for the neutral charge state, about 0.11 eV below E_F . Thus, for higher precision we here subtract this 0.11 eV from E_F to obtain the defect level. This would reduce the a_1 defect levels in Table II

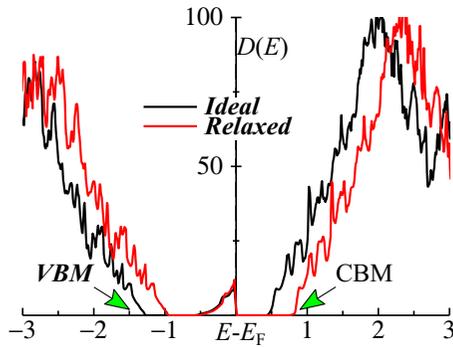


FIG. 9. Density-of-states of a 128-atom supercell of GaAs with a single As antisite for an unrelaxed lattice (black) and for a lattice whose nearest neighbors were relaxed (red). Energy zero corresponds to the Fermi level in both cases, and sits at the top of the (doubly occupied) midgap EL2 level. It has a dispersion of about 0.3 eV in the 128-atom cell.

by 0.11 eV. The VBM and CBM are inferred from the energies where the DOS touches zero (see labels in Fig. 9), so the EL2 relative to either can be computed. The results are displayed in Table III.

As Table III shows, the (+/0) level of EL2 is predicted to be at $E_v + 0.83$ eV and $E_c - 0.85$ eV. It is slightly too far from both the VBM and the CBM, compared to experiment, because the QSGW gap is too large. It is not *a priori* obvious how much this level will shift if QSGW were high enough fidelity to yield the experimental gap, e.g., by adding ladder diagrams to W . It is possible, in principle, to do this by carrying out the calculation with ladders in W , but here we take the hybrid 80% QSGW + 20% LDA approach as a simple alternative. The result is shown in Table III. The hybrid approach reduces the level by about 0.1 eV, while the gap itself is reduced by ~ 0.25 eV. This is consistent with the EL2 being comprised of roughly equal measures of the host valence band and conduction band character and its position almost exactly in the middle of the gap. Finally, the predicted (+/0) EL2 energy ($E_v + 0.74$ eV) is in excellent agreement with the observed value $E_v + 0.75$ eV, while the band gap (1.42 eV) is also close to the room-temperature experimental gap. (Ideally, the QSGW gap should be ~ 1.5 eV, the zero-temperature gap of GaAs. Indeed, it does come out very close to 1.5 eV when a fully converged basis is used, whether ladders are added or the hybrid- Σ approach is used).

TABLE III. Band center of neutral EL2 defect, in eV, relative to the valence band maximum, embedded in a 128-atom cell. A reasonably well converged basis was used, with spin orbit coupling included. Two calculations are shown: The first assuming no lattice relaxation and the second including it, as described in the text. Also shown is the band gap.

Structure	QSGW		80% QSGW + 20% LDA	
	EL2	gap	EL2	gap
Unrelaxed	1.13	1.69	1.02	1.44
Relaxed	0.83	1.68	0.74	1.42

We can also estimate the (2 + /+) transition level. The latter is experimentally found at 0.52–0.54 eV or about 0.23 eV lower than the (+/0) level. Following the approach of Bachelet *et al.* [27], we can use the difference between our $q = 2+$ one-electron level and $q = 0$ level to estimate the transition levels. Using Janak's theorem, $\epsilon_i = \partial E(n_i)/\partial n_i$, and assuming a linear dependence $\epsilon_i(n_i) = \epsilon(0) + U n_i$, we obtain $E(n_i) = \epsilon_i(0)n_i + U n_i^2/2$ or $\epsilon(+/0) = E(2) - E(1) = \epsilon_i(0) + \frac{3}{2}U$ and $\epsilon(2 + /+) = E(1) - E(0) = \epsilon_i(0) + \frac{1}{2}U$. From Table II, our value of $U \approx 0.2 - 0.3$ eV including the charge-background correction, in excellent agreement with the experimental value of 0.23 eV. Thus, if with the large basis set we obtain the (+/0) level at 0.74 eV, our value for the (2 + /+) level will also be close.

Even with a slightly less complete basis set, including $spdfspd$ and Ga-3d local orbitals or $spdsp+$ local orbitals, but omitting the floating orbitals, we found that the defect band top in a 64-atom cell and using 64₈ cut-and-paste approach lies at 0.78 eV above the VBM when the latter is corrected by SOC, and with a host gap of 1.7 eV (1.5 eV for the $spdsp$ basis set). Given that the center of gravity of the defect band DOS lies slightly lower, this would become 0.67 eV. Even in this calculation, however, the CBM in the defect cell is ~ 1.35 eV, indicating that the t_2 resonance in the conduction band affects the CBM. This effect is reduced only by going to even larger supercells. Using the $spdsp$ basis set, we have also calculated the defect level in $q = 0$ and $q = +2$ charge states and find them to be 0.81 eV and 0.48 eV (including the charge-background correction and the spin-orbit coupling gap correction). These did not yet include the 0.8Σ correction but are consistent with the large basis set and show that both the +/0 and 2 + /+ transitions are in good agreement with experiment.

We finally note that while these levels are indicated as +/0 and 2 + /+, the experiments are more directly related to the frozen geometry one-electron excitation levels we calculate here in, respectively, the neutral and 2+ geometries because they correspond to optical excitations activating the EPR singly occupied $q = +1$ state and not thermodynamic transition levels. To calculate the thermodynamic transition levels, one can follow the approach suggested by Rinke *et al.* [5]. For example, the thermodynamic transition level

$$\begin{aligned}
 \epsilon(+/0) &= E_f(0, 0) - E_f(+, +) - \epsilon_F, \\
 &= [E_f(0, 0) - E_f(+, 0)] \\
 &\quad + [E_f(+, 0) - E_f(+, +)] - \epsilon_F, \\
 &= \epsilon_d - \epsilon_F + E_{\text{relax}}(0 \rightarrow +). \tag{11}
 \end{aligned}$$

Here $E_f(q, q')$ indicates the energy of formation of the defect in charge state q in the relaxed geometry of charge state q' . The first bracket in the second line is frozen geometry excitation energy from the neutral to the $q = +1$ charge state, calculated in the GW approach. If we pick the chemical potential a ϵ_F at the VBM, then $\epsilon_d - \epsilon_v$ is just the one-electron defect level relative to the VBM. The second bracket is the relaxation energy of the $q = +1$ charge state from the initial $q = 0$ geometry to its own equilibrium geometry. This amounts to about 0.08 eV.

IV. CONCLUSION

We have shown that the GW self-energy matrix represented in a real-space basis set is short-ranged, and the contribution to the quasiparticle energy correction to DFT eigenvalues is significant only for the first few nearest neighbors of a specific atom. We introduced a cut-and-paste method for defect calculations at the GW level that exploits this property. We demonstrated the method using a well-known single point defect, namely, the As_{Ga} antisite in a GaAs crystal. The main correction to the defect band structure compared to LDA in our method is incorporating the host perfect crystal gap correction via the perfect crystal self-energy, which requires a trivial cost because it just amounts to a relabeling of the self-energy matrix according to the supercell description of the atomic sites. After detailed examination, we conclude that an eight-atom cell is sufficient to extract the defect atom's self-energy which is then used to replace the defect atom self-energy in the final supercell. We observe almost perfect agreement between the fully self-consistent QSGW defect bands and the cut-and-paste method, in terms of the valence bands, the a_1 defect-level position, and the defect band dispersion. There is a small disagreement between our method and the full QSGW calculations, in terms of unoccupied levels. This is caused mainly by the range cutoff d_{max} applied to Σ , which slightly reduces the gap from its converged QSGW value. The main advantage of the method is that it allows us to obtain GW -level accuracy results for large defect supercells at essentially the cost of an LDA calculation for the latter. This allowed us to carefully monitor the defect band dispersion and identify the dilute limit isolated defect level more precisely. We found good agreement for the defect-level positions with

previous studies of this system, although the previous studies were not at the GW level. Inspecting the band structures of the defect system in detail allowed us to identify not only the obvious defect level in the gap of a_1 symmetry but also the excited defect t_2 symmetry resonance and provides accurate information on the optical transition between these levels, which has previously been recognized as an important step in activating a metastable state of this defect. Our calculation also agrees with previous work in the change in defect level as a function of charge states. Overall, the cut-and-paste method significantly reduces the computational cost of GW -level calculations, with a small loss in accuracy. Finally, to establish that QSGW is able to predict defect levels with a fidelity comparable to its ability to predict energy bands of bulk materials, we benchmark the embedding approach against the experimentally measured neutral EL2 level. We show the discrepancies with experiments are small, and closely track the known discrepancies for weakly correlated periodic systems.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy Basic Energy Sciences (DOE-BES) under Grant No. DE-SC0008933. The calculations were performed on the High Performance Computing Resource in the Core Facility of Advanced Research Computing at Case Western Reserve University. M.v.S. was supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Division of Chemical Sciences, under Contract No. DE-AC36-08GO28308.

-
- [1] L. Hedin, *Phys. Rev.* **139**, A796 (1965).
 - [2] L. Hedin and S. Lundqvist, in *Solid State Physics, Advanced in Research and Applications*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich, Vol. 23 (Academic Press, New York, 1969), pp. 1–181.
 - [3] M. van Schilfgaarde, T. Kotani, and S. Faleev, *Phys. Rev. Lett.* **96**, 226402 (2006).
 - [4] C. Freysoldt, B. Grabowski, T. Hickel, J. Neugebauer, G. Kresse, A. Janotti, and C. G. Van de Walle, *Rev. Mod. Phys.* **86**, 253 (2014).
 - [5] P. Rinke, A. Janotti, M. Scheffler, and C. G. Van de Walle, *Phys. Rev. Lett.* **102**, 026402 (2009).
 - [6] S. Lany and A. Zunger, *Phys. Rev. B* **81**, 113201 (2010).
 - [7] W. Chen and A. Pasquarello, *J. Phys.: Condens. Matter* **27**, 133202 (2015).
 - [8] W. Chen and A. Pasquarello, *Phys. Rev. B* **96**, 020101(R) (2017).
 - [9] T. Kotani, M. van Schilfgaarde, and S. V. Faleev, *Phys. Rev. B* **76**, 165106 (2007).
 - [10] N. E. Christensen, *Phys. Rev. B* **30**, 5753 (1984).
 - [11] T. R. Paudel and W. R. L. Lambrecht, *Phys. Rev. B* **77**, 205202 (2008).
 - [12] A. Boonchun and W. R. L. Lambrecht, *Phys. Status Solidi B* **248**, 1043 (2011).
 - [13] D. Skachkov, A. Punya Jaroenjittichai, L.-y. Huang, and W. R. L. Lambrecht, *Phys. Rev. B* **93**, 155202 (2016).
 - [14] S. Lany, H. Raebiger, and A. Zunger, *Phys. Rev. B* **77**, 241201(R) (2008).
 - [15] D. Segev, A. Janotti, and C. G. Van de Walle, *Phys. Rev. B* **75**, 035201 (2007).
 - [16] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207 (2003).
 - [17] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **124**, 219906 (2006).
 - [18] J. P. Perdew, M. Ernzerhof, and K. Burke, *J. Chem. Phys.* **105**, 9982 (1996).
 - [19] S. Lany and A. Zunger, *Phys. Rev. B* **80**, 085202 (2009).
 - [20] S. Lany and A. Zunger, *Phys. Rev. B* **81**, 205209 (2010).
 - [21] H. Ma, N. Sheng, M. Govoni, and G. Galli, *J. Chem. Theory Comput.* **17**, 2116 (2021).
 - [22] F. Aryasetiawan and O. Gunnarsson, *Phys. Rev. B* **49**, 16214 (1994).
 - [23] T. Kotani, *J. Phys. Soc. Jpn.* **83**, 094711 (2014).
 - [24] C. Friedrich, S. Blügel, and A. Schindlmayr, *Phys. Rev. B* **81**, 125102 (2010).

- [25] 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).
- [26] <https://www.questaal.org/>. Our *GW* implementation was adapted from the original ECALJ package now at <https://github.com/tkotani/ecalj/>.
- [27] G. B. Bachelet, M. Schlüter, and G. A. Baraff, *Phys. Rev. B* **27**, 2545 (1983).
- [28] J. Dabrowski and M. Scheffler, *Phys. Rev. B* **40**, 10391 (1989).
- [29] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.105.205136> for additional test results with other parameter choices and a flow chart of the scheme.
- [30] E. R. Weber, H. Ennen, U. Kaufmann, J. Windscheif, J. Schneider, and T. Wosinski, *J. Appl. Phys.* **53**, 6140 (1982).
- [31] J. Lagowski, D. G. Lin, T. Chen, M. Skowronski, and H. C. Gatos, *Appl. Phys. Lett.* **47**, 929 (1985).
- [32] P. Omling, E. R. Weber, and L. Samuelson, *Phys. Rev. B* **33**, 5880 (1986).
- [33] L. Samuelson and P. Omling, *Phys. Rev. B* **34**, 5603 (1986).
- [34] A. N. Chantis, M. van Schilfgaarde, and T. Kotani, *Phys. Rev. Lett.* **96**, 086405 (2006).
- [35] C. Bhandari, M. van Schilfgaarde, T. Kotani, and W. R. L. Lambrecht, *Phys. Rev. Materials* **2**, 013807 (2018).
- [36] M. Rohlfing and S. G. Louie, *Phys. Rev. B* **62**, 4927 (2000).
- [37] S. Albrecht, L. Reining, R. DelSole, and G. Onida, *Phys. Rev. Lett.* **80**, 4510 (1998).
- [38] B. Cunningham, M. Gruening, D. Pashov, and M. van Schilfgaarde, [arXiv:2106.05759](https://arxiv.org/abs/2106.05759).
- [39] D. Deguchi, K. Sato, H. Kino, and T. Kotani, *Jpn. J. Appl. Phys.* **55**, 051201 (2016).