The carbon state in dilute germanium carbides

Istvan A. Gulyas, Chad A. Stephenson, Qian Meng, Seth R. Bank, and Mark A. Wistey

View Online   Export Citation   CrossMark
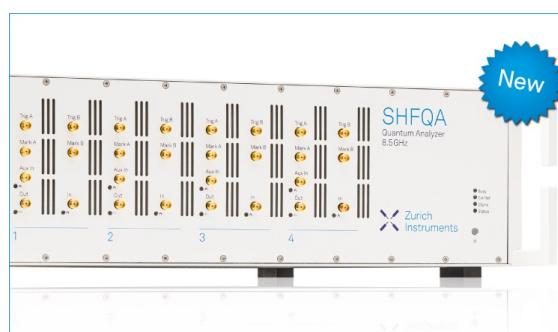
The carbon state in dilute germanium carbides

View Online   Export Citation   CrossMark

Istvan A. Gulyas,[1] Chad A. Stephenson,[2,b] Qian Meng,[3] Seth R. Bank,[3] and Mark A. Wistey[1,a]

AFFILIATIONS

[1] Department of Physics, Texas State University, San Marcos, Texas 78666, USA
[2] Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556, USA
[3] Electrical & Computer Engineering Department, University of Texas at Austin, Austin, Texas 78705, USA

Note: This paper is part of the Special Topic on Highly Mismatched Semiconductors Alloys: from Atoms to Devices.
[a] Author to whom correspondence should be addressed: mwistey@txstate.edu
[b] Present address: Sandia National Laboratory, Albuquerque, NM 87185.

ABSTRACT

Conduction and valence band states for the highly mismatched alloy (HMA) Ge:C are projected onto Ge crystal states, Ge vacancy states, and Ge/C atomic orbitals, revealing that substitutional carbon not only creates a direct bandgap but also the new conduction band is optically active. Overlap integrals of the new Ge:C conduction band state with states from unperturbed Ge show that the new band cannot be attributed to any single Ge band but is a mixture of multiple Ge states. The Ge Γ conduction band valley state plays the largest single role, but L and X valley states collectively contribute a larger share than Γ due to the multiplicity of degenerate states. C sites structurally resemble uncharged vacancies in the Ge lattice, similar to Hjalmarson's model for other HMAs. C also perturbs the entire Ge band structure even at the deepest crystal core energy levels, particularly if staggered supercells are used to mimic a disordered alloy. Projection onto atomic sites shows a relatively weak localization compared with other HMAs, but it does show a strong anisotropy in probability distribution. L-valley conduction band states in Ge contribute to the conduction band minimum in Ge:C, but the optical transition strength in Ge: C remains within a factor of 2 of the direct gap transition in Ge.

I. INTRODUCTION

Dilute germanium carbides (Ge:C) are a highly promising candidate for direct bandgap tunneling[1] and photonic[2,3] devices on silicon substrates[4] with growth techniques that avoid undesirable carbon–carbon bonds during growth.[5–8] Highly mismatched alloys (HMAs) such as Ge:C, GaAs$_{1-x}$N$_x$, and GaAs$_{1-x}$B$_x$ exhibit properties beyond the ranges predicted by the virtual crystal approximation.[7,9,10] This opens new wavelengths for lasers,[11,12] detectors, and solar cells[13] while remaining lattice matched or at least compatible with Si, Ge, GaAs, or InP. Similarly, independent control of material properties could greatly improve steep switching in tunneling field effect transistors. Most semiconductors are constrained by tight coupling between material properties such as bandgap, lattice constant, and effective mass. But HMAs provide additional degrees of freedom even at small compositions of the mismatched atom, typically <3 at. %. For example, the addition of only 2% nitrogen to InGaAs reduces the bandgap by almost 200 meV while simultaneously reducing strain on GaAs substrates.[14] At the other extreme, adding small amounts of B to GaAs appears to leave critical points

of the band structure almost unchanged while reducing the lattice constant considerably.[15]

The mismatch in HMAs comes from the alloying of elements with the same valence but very different sizes, bond angle, and/or electronegativity. Such constraints typically reduce solubility limits to much less than 1% of the mismatched atom. This makes most HMAs difficult to synthesize except in very dilute quantities. Even kinetically limited techniques such as conventional molecular beam epitaxy may be unable to reach compositions approaching 1%.[16] The mismatch also makes HMAs a challenge to simulate numerically because the approximations that are frequently used for one atom species may be invalid for the other. For example, computational models of BN can ignore core electrons (none) and often omit spin–orbit coupling, but the "hard" or rapidly varying potential near the N nucleus requires a large basis set of plane waves to accurately represent wavefunctions. On the other hand, "soft" InAs can be modeled with a smaller basis set but inner d electrons and a relativistic Hamiltonian play a significant role. Strictly speaking, an accurate computational model of an HMA such as InAs$_{1-x}$N$_x$,

therefore, requires the combination of large supercells, many electrons per atom including inner electrons, a large basis set of plane waves, and spin–orbit coupling, resulting in an unfeasible computational demand.

To make such calculations tractable, GW many-body perturbation techniques have been used to calculate the energies of defects in semiconductors[17] but large supercells were previously considered too computationally expensive for such techniques.[18] Unfortunately, the small bandgap of Ge leads to a metallic band structure when using supercells as small as 64 atoms (1.6 at. % C). Even larger supercells are necessary if simulating the smaller carbon concentrations corresponding to bandgaps used for photonic integrated circuits or datacom, necessitating 128 atom supercells (0.78 at. % C).

Previous reports explored the band structures of Ge:C and other HMAs. In dilute nitrides such as $GaInAs_{1-y}N_y$, band anticrossing (BAC) provides a satisfying, if somewhat simplified, prediction of band structures, which are readily modeled using $k \cdot p$ perturbation methods with a single additional band. Nitrogen creates a state within the conduction band that behaves atypically with pressure, as if it were pinned to the vacuum level rather than to the conduction band edge. This behavior is usually described using the deep trap model of Hjalmarson and Dow, in which the defect state is dominated by dangling bonds in the host.[19,20] In this context, "deep" is not "deep within the bandgap" but it refers to states whose origins are very deep valence atomic states or other states whose behavior is tied to the vacuum level, rather than the conduction or valence band maxima (VBM).

In light of recent reports that the lowest conduction band in Ge:C is optically inactive,[21] i.e., a pseudo-direct bandgap,[22] and after establishing suitable convergence conditions for accurate modeling of Ge:C,[23] we investigated the nature of the carbon states in Ge:C. In particular, we examined whether the electronic and optical properties of Ge:C and their pressure dependence (or the lack thereof)[20] can be explained using Hjalmarson's model. We applied computational techniques with higher accuracy than were previously available. Although BAC correctly predicts a splitting of the conduction band at $k = 0$, it fails to describe the band structure at higher values of k, in contrast to many other HMAs such as $GaAs_{1-x}N_x$.

In this report, we use dopant notation (Ge:C) since the fraction of C, 0.78%, is quite small for an alloy. However, its effects are far more pronounced than typical dopants.

II. COMPUTATIONAL METHODS

The Vienna *ab initio* Simulation Package (VASP)[24–27] was used to perform density functional theory (DFT) based simulations of a 128-site diamond lattice supercell of Ge, Ge:C, or Ge with a vacancy (v_{Ge}). For v_{Ge}, one Ge atom was removed; for Ge:C, it was replaced with C. The projector-augmented wave (PAW) core electron method was used with the generalized gradient approximation (GGA) and the Perdew–Burke–Ernzerhof (PBE) functional.[28–31] PBE functionals are well known to underestimate bandgaps but all-electron models are prohibitive due to the sheer numbers of atoms (128) and electrons (32 per Ge atom). Given N total electrons, the complexity varies from $(N^2 \ln N)$ for FFT-limited techniques to $O(N^3)$ if exactly diagonalizing the Hamiltonian. As a

compromise, the Heyd–Scuseria–Ernzerhof (HSE06) hybrid functional was used instead of PBE after the first ionic relaxation.[32] The results from HSE06 and a similar functional, HSEsol, have been reported to be very similar.[33] Due to computational limits and the large supercells used here, the inclusion of PBE0 hybrid functionals, spin–orbit coupling (SOC), and d electrons was generally impractical, although a few trial results are given below. Although SOC and d electrons affect the valence band,[34] they were previously found to induce relatively minor changes in the Ge conduction band structure, in terms of both calculations[5] and theory.[35] Further simulation details can be found in Ref. 23.

Unless otherwise noted, the following parameters were used for the calculations presented here. The computational lattice constant was calculated by rigidly varying the supercell lattice vectors, relaxing the ion positions within each new volume, and fitting the resulting set of system energies to the Birch–Murnaghan equation of state. These computational lattice constants (Table I) were then used for subsequent calculations without attempting to force a fit to experimental parameters such as direct or indirect bandgaps. PAW PBE potentials with outer radii of 2.342 Å and 2.266 Å were used for Ge and C, respectively. Projection operators were evaluated in reciprocal space. The HSE screening parameter was slightly modified from 0.20 (HSE06) to 0.18 for comparison with previous work.[5]

In addition to hybrid functionals, we increased the number of k points and plane waves beyond the default values for typical simulations, in order to more accurately capture short-range electronic structures. Specifically, we used a Γ-centered $2 \times 2 \times 2$ mesh of k points, which would be comparable with $8 \times 8 \times 8$ in the two-atom primitive cell. An energy cutoff of 600 eV was used for plane wave basis sets as this was found to be well converged in system energy.[23] To accurately capture higher conduction band states, free carrier absorption, and effective masses, 784 bands were included in the calculations, many of which were degenerate, particularly at higher levels. For comparison, the 256th band was the last filled valence band, and the rest were conduction bands; adding additional bands better captured not only conduction band curvatures (effective masses) but also the character of upper conduction band states, and, therefore, higher optical transitions such as free carrier absorption (FCA).

VASP normalizes its wavefunctions using an overlap operator rather than directly setting the norm $\psi_n^* | \psi_n = 1$. The difference is typically relatively small. For example, we found that a two-atom GaAs cell with 56 bands showed a maximum norm $\psi_n^* | \psi_n$ of 1.45, a minimum of 0.894, and a standard deviation of 0.099. Similarly, for 128-atom Ge:C supercells with 336 bands presented here, the

TABLE I. Relaxed (i.e., computational) lattice constants, band energies, and projections on to s orbitals for the supercells of Ge_{128}, Ge with a single vacancy (v_{Ge} or $Ge_{127}v_1$), and $Ge_{127}C_1$. %s is the ratio of s to the sum of all s, p, and d orbitals in the projection.

Material	Lattice const. (Å)	$E_{G,\Gamma}$ (eV)	$E_{G,L}$ (eV)	%s CB_Γ	%s CB_Γ on C atom
Ge_{128}	5.673	0.788	0.804	100	N/A
v_{Ge} ($Ge_{127}v_1$)	5.667	1.093	0.685	90	N/A
Ge:C	5.655	0.434	0.825	80	100

range of norms was 1.04–1.37 with a standard deviation of 0.083. However, in order to accurately compare even small differences in either overlaps (inner products) or optical (momentum) matrix elements between two states, the wavefunctions of the two states were normalized using the components of the wavefunction as follows. Given a plane wave basis set at a given k and truncated at $|\mathbf{G} + \mathbf{k}| \leq \mathbf{G}_{cut} = \sqrt{2mE_{max}}/\hbar$,

$$\psi_k(\vec{r}) = \sum_G a_{n,k} e^{-i(\vec{G}+\vec{k})\cdot\vec{r}}. \tag{1}$$

The inner product or overlap integral between states ψ_i and ψ_f is

$$\left\langle \psi_{i,k}^* | \psi_{f,k} \right\rangle = \sum_G a_{n,k} b_{n,k}, \tag{2}$$

where a and b are the coefficients of the plane waves forming the initial and final states, respectively. These plane wave coefficients were extracted from the WAVECAR file using WaveTrans.[36] If ψ_i and ψ_f used different basis sets in k, then the overlap integral was instead calculated in real space over the supercell as follows, using PyVaspwfc to extract wavefunctions,[37]

$$\left\langle \psi_{i,k}^* | \psi_{f,k} \right\rangle = \iiint \psi_{i,k}^*(\vec{r}) \psi_{f,k}(\vec{r}) d^3r. \tag{3}$$

To account for the arbitrary temporal phase of the wavefunctions, $b_{n,k}$ (or $\psi_{f,k}$) was multiplied by a phase term $e^{i\varphi}$, which was varied to maximize the integral. For the calculations reported here, the wavefunctions or plane wave coefficients returned by VASP at each value of $\vec{k}$ were independently rescaled for a norm of 1, i.e.,

$$\left\langle \psi_{i,k}^* | \psi_{i,k} \right\rangle = 1. \tag{4}$$

Equation (3) was used to project one state onto another, such as when quantifying the similarity between the C states in Ge:C and the original unmodified conduction band states in Ge. Similarly, the relative transition strength varies as the square of the momentum matrix element $P_{i,f}$ of optical transitions between two states $\psi_{i,k}$ and $\psi_{f,k}$, determined from the momentum operator and Fermi's golden rule,

$$P_{i,f} = \left\langle \psi_{i,k}^*(\vec{r}) \,|\, H'(\vec{r})| \, \psi_{f,k}(\vec{r}) \right\rangle = \left\langle \psi_{i,k}^*(\vec{r}) \,|\, -i\hbar\hat{e}\cdot\nabla \,|\, \psi_{f,k}(\vec{r}) \right\rangle, \tag{5}$$

where $\hat{e}$ is the unit vector of the optical electric field, which we assume from here on to be polarized such that the dot product is maximized. Dipole matrix elements from VASP were scaled using the same renormalization factors as in Eq. (4).

In this work, the HSE screening parameter was slightly modified from 0.20 (HSE06) to 0.18 for comparison with previous work[5] but computational lattice constants were not adjusted. Forcing a smaller lattice constant to fit both direct and indirect bandgaps would not directly affect the main points of this work, although it could influence whether a given alloy and strain are

direct or indirect if the difference in energies is small. Also, in the absence of both spin–orbit coupling and strain, the three valence bands are all degenerate at Γ.

When the crystal primitive cell is repeated to form a supercell, the first Brillouin zone gets folded onto itself. For the $4 \times 4 \times 4$ supercells used here, the L and X edges of the Brillouin zone fold onto Γ, and then the Brillouin zone gets folded yet again. Thus, the energy eigenvalues at "$k = 0$" actually contain the union of the band energies for (in primitive cell Cartesian coordinates) $k = (0, 0, 0), \pi/a(\pm 1, 0, 0), \pi/a(\pm 1, \pm 1, \pm 1), \pi/2a(\pm 1, 0, 0), \pi/2a(\pm 1, \pm 1, \pm 1)$, etc., where a is the lattice constant, making interpretation nontrivial. To identify the equivalent k in the primitive cell Brillouin zone, either the bands were unfolded using either BandUP,[37–39] vasp_unfold,[40] or PyVaspwfc[37] or the character of the band was determined manually from the projection onto the s and p orbitals, which is also the method used by vasp_unfold. Because each state is divided among 128 atoms, to reduce rounding errors when projecting onto atomic orbitals, VASP was modified to produce six digits of precision in its PROCAR output files. Also, due to the periodic boundary conditions, it does not matter which Ge atom is replaced by C since the overall periodicity and resulting band structures would be identical.

III. RESULTS AND DISCUSSION

A. Orbital character

Adding dilute C to Ge has previously been shown to decrease the bandgap at Γ, as in dilute nitrides such as $GaAs_{1-x}N_x$. The bottom conduction band (CB) is split into three bands at Γ, identified later as E^+, E_2, and E^-, along with an *increase* in the conduction effective mass; only two of these are predicted by the BAC model. Also, the effects of C on the band structure are not well captured by BAC away from $k = 0$ with a plateau in E^- toward L and a rise toward X. What are the band and bond origins of the carbon state in Ge, and why does it show such asymmetry in the band structure?

Figure 1 and Table I show the projection onto atomic orbitals of the Γ states at the valence band (VB) maximum, which is triply
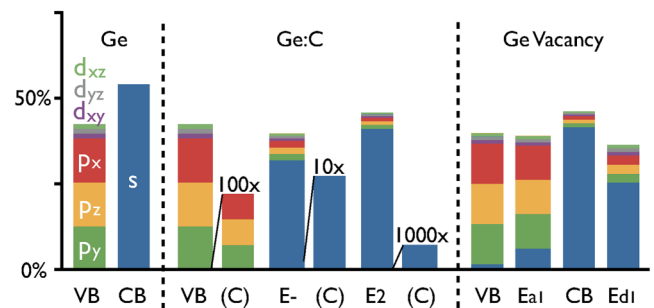
FIG. 1. Atomic orbital character of states near the bandgap at Γ. The projections onto orbitals are summed over all atoms, except that projections onto the C ion alone are also shown for Ge:C. Degenerate bands are averaged. Totals are less than 100% because only outer orbitals and contributions within the Wigner–Seitz radius of each atom are included.

degenerate, as well as the two lowest CB minima. 128-atom super-cells of Ge, $Ge_{127}C_1$, and $Ge_{127}v_1$ (single Ge vacancy) were used with a $2 \times 2 \times 2$ mesh of k points. Both of the lowest two CBs in Ge:C are predominantly s-like in character over most atoms, like the unperturbed Ge CB, and the VBs likewise retain their predominantly p-like character. This suggests that both E_2 and E^- CBs in Ge:C will be optically active for band-to-band transitions. The similar s-like nature of the split E^+/E_2 CBs does match the premise and predictions of the band anticrossing model at Γ, although we shall show later that the first-order BAC model rapidly fails to explain E_2 or band structures away from k = 0. Adding a vacancy in Ge (v_{Ge} or $Ge_{127}v_1$) introduces a new, empty p-like state 0.68 eV above the VB edge and a mostly s-like state above the CB edge, tentatively identified as acceptor-like and donor-like states, E_{a1} and E_{d1}, respectively. The Ge:C E^+/E^- bands appear to have very different characters from the vacancy state, though both C and vacancies are treated in Hjalmarson's model as "deep" states.[19] However, we shall see later that other similarities do exist. Also, the E^- band is overrepresented by the C atom, while the E^+ band has almost no overlap with the C atom. This qualitatively agrees with a simple perturbation model between two similar states |a> and |b>

producing new states |a+b> and |a−b>. We note in passing that the E_{a1} state is triply degenerate, like the VB.

B. Comparisons with bulk Ge and Ge vacancy states

Although the projection onto atomic orbitals provides a qualitative measure of localization on the C atom, it may lose important quantitative information about the character of the state within the band, e.g., optical or transport properties. To address this, we calculated the inner product of the E^- state (and the E^+ state) with unperturbed Ge wavefunctions that were simulated using the same basis set and conditions as Ge:C. This allows the identification of which bands in the Ge crystal the alloy mixed with to create the carbon state(s).

In order to identify how the C states gain their character, including pressure dependence, states near the Ge:C band edge were compared with those from either pure Ge or Ge with a single vacancy (v_{Ge}). As shown in Fig. 2, the Ge:C valence bands and *upper* conduction band E+ are seen to share character with the corresponding bands in Ge. But the lower conduction band, E^-, is not well modeled by any single Ge band. Rather, E^- is a mix of several
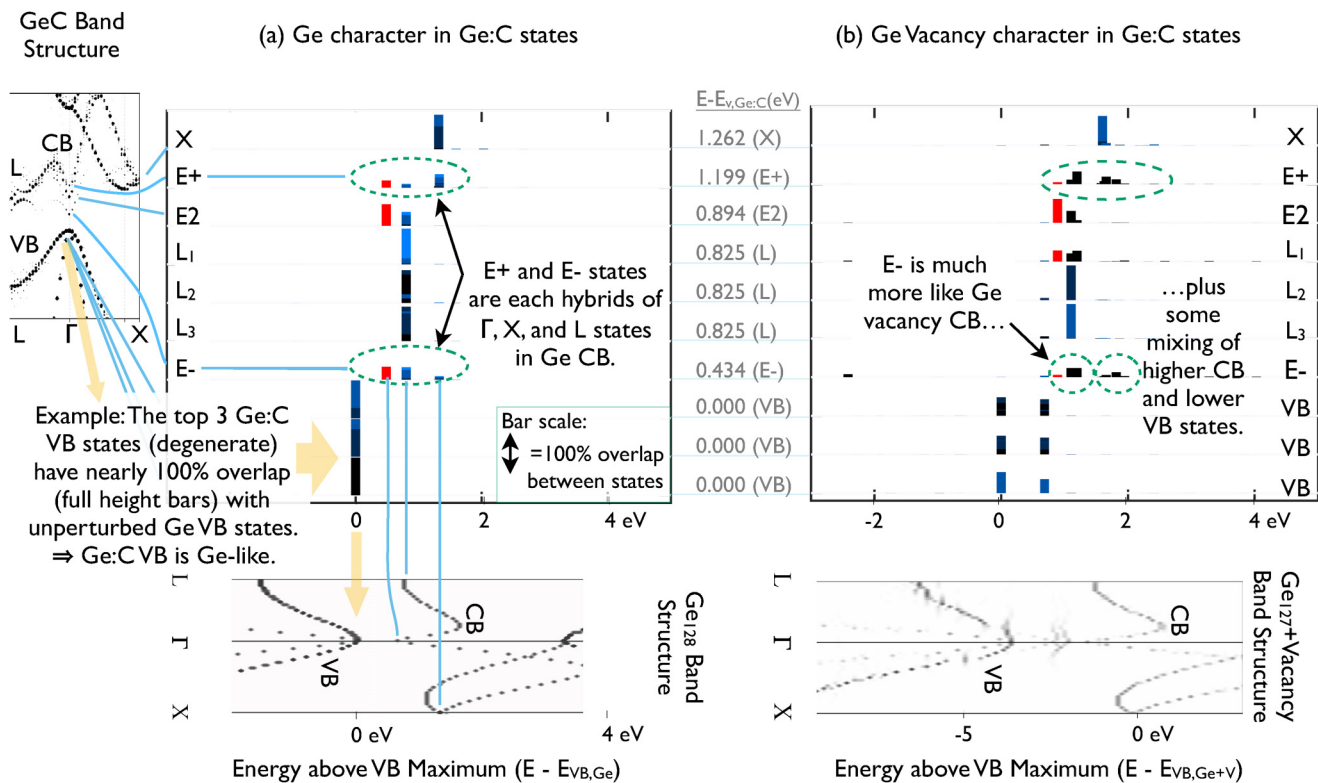
FIG. 2. Projection of several Γ and Γ-folded Ge:C states onto states from either (a) pure Ge_{128} or (b) Ge with a single vacancy. The length of each bar represents the squared inner product between the given Ge:C state and the Ge or v_{Ge} state at the given energy. For clarity, bars from closely spaced bands are stacked in different shades of blue, rather than overlapping, and GeC bands are offset in equal vertical steps, labeled as energy (eV) above Ge:C VBM. Also, overlaps with Ge and v_{Ge} Γ CBM are drawn in red and offset left to avoid overlap from near-degeneracy with L. Straight blue and green lines are guides to the corresponding points in the respective band structures.

TABLE II. Projection of the Ge:C E⁻ CB state onto various states in unperturbed Ge. To reduce numerical noise, only states with overlaps >0.01 are included. Degenerate bands are summed.

Ge state	CB Γ	CB L	CB X	All other bands
Bandgap (eV)	0.788	0.804	1.330	
Degeneracy	1	4	6	
ΔE/strain (eV/%)	−0.323	−0.113	0.034	
Overlap with E⁻	0.34	0.37	0.10	0.00
% of RMS total	42%	46%	12%	0%
Weighted avg.		−0.183 eV/%		

Ge bands at Γ with a significant projection onto the defect state, as circled in Fig. 2(b). The presence of multiple bars per row (corresponding to a single Ge:C state) indicates mixing from different states across a range of energies. Higher CBs also have some low-VB character.

Table II shows a breakdown of the E⁻ state projected onto states in unperturbed Ge, as in Fig. 2. Only states with an overlap of >0.03 are included. The overlaps show that the E⁻ band is a mixed state of approximately 42% Ge Γ CB, 46% Ge L CB, and 12% Ge X CB. The pressure dependence of these Ge states will be discussed in Sec. IV D. Additional similarities between the C "defect" and a Ge vacancy are shown in the filled-state charge density, as shown in Fig. 3. In v_{Ge} and Ge:C, the defects visibly affect longer range charge densities, at least to the third-nearest neighboring atoms. Surprisingly, although v_{Ge} and Ge:C show nearly identical charge distributions away from the defect, Ge:C shows significant charge strongly localized on the C atom. Since this is a filled-state plot, it shows that the charge on C must arise from valence states (i.e., filled states). Additionally, since the charge is strongly localized, it is also strongly bonded, arising from states at low energies deep within the valence bands.

C. Γ character of the E⁻ conduction band

The E⁻ band at Γ was reported by Kirwan *et al.* to vary weakly with pressure, which was interpreted to mean that L states dominated the E⁻ band. This would lead to a pseudo-direct bandgap[22] (not to be confused with a nearly direct bandgap) in which the lowest conduction band states had the wrong symmetry for strong optical transitions to or from the valence band. As mentioned above, optical transitions impose a momentum operator in the inner product between the initial and the final states, so a symmetric final state becomes antisymmetric in the integral, and vice versa. Therefore, the strongest transitions are those from symmetric to antisymmetric states, or vice versa, such as between the *s* and *p* states. To examine whether this independence was due to L-valley states or deep *s*-like states instead, we extracted optical transition momentum matrix elements for transitions from the valence bands to the E⁻ band. The projection onto atomic orbitals in Fig. 1 verified that the valence bands were still 90% *p* orbitals, which are antisymmetric with respect to the atom cores. The electron states at L are likewise antisymmetric due to the Bloch waves at the edge of the Brillouin zone, so the addition of L states at Γ would tend to reduce the strength of optical transitions.

Instead, we found that the momentum matrix element (the strength of the optical transition, $|P_{cv}|^2$) from the VB maximum to CB minimum (E⁻) in Ge:C was within 50% of that in Ge. Also, the VB-E⁺ transition was within 26% of Ge but all other nearby transitions were two orders of magnitude smaller. This supports the conclusion that both E⁻ and E⁺ CBs retain a strong component of *s*-like symmetry. Optical transitions will be reported in more detail elsewhere. The reason for the strong optical transition is revealed by plotting the charge densities for a valence band state alongside the E⁻ state (PARCHG files), plotted in Fig. 4 using VESTA software.[41] [Simulation conditions: 128 atoms, 336 bands, single k-point (Γ), cutoff energy 400 eV.] VB states are overwhelmingly mirror-symmetric along bonds. The E⁻ state is spread among not only many atoms but predominantly only on one atom at each end of bonds, leading to strong odd symmetry along the bond. Therefore, optical transitions are also strong. The symmetry and strong optical transitions to the E⁻ band also mean that the reported pressure independence comes from a source other than indirect L-valley states, again consistent with band anticrossing from the defect state in Hjalmarson's model at Γ.
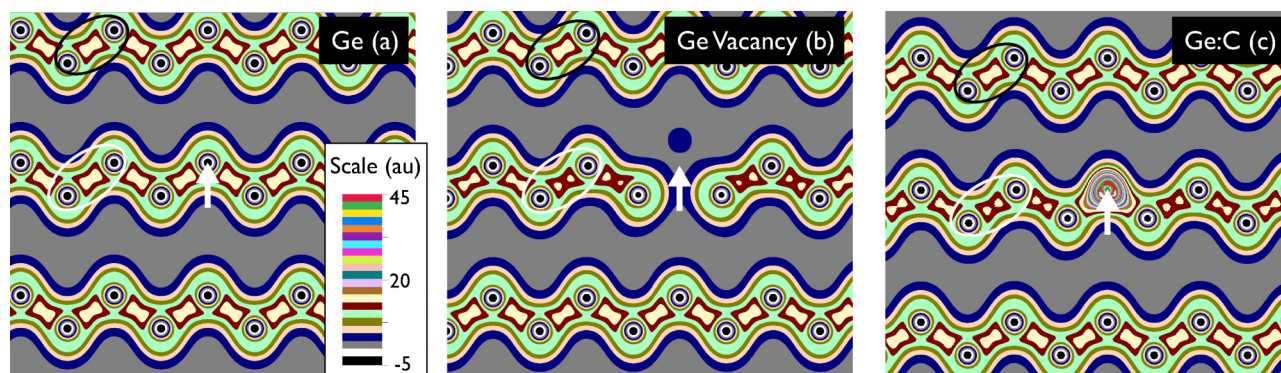
FIG. 3. Filled-state charge density pseudocolor plots along a {110} plane. (a) Unperturbed Ge, (b) Ge with a single vacancy, and (c) Ge:C. Equivalent positions are marked with white arrows. Although a C atom attracts charge and the vacancy repels it, both types of defect reduce the charge in bonds as far away as third-nearest neighbors along ⟨110⟩ (white ovals) compared with other directions (black ovals) or any bond in unperturbed Ge. Color scales are identical.
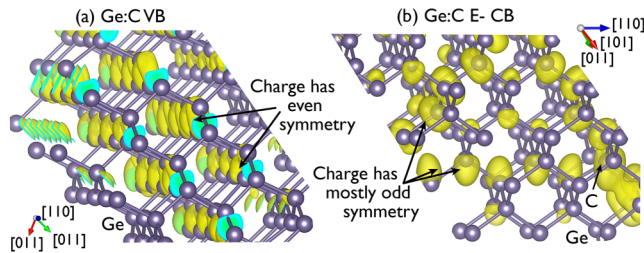
FIG. 4. Ge:C partial charge density isosurface (yellow) from (a) one of the valence band states and (b) the E^- state, which is not only delocalized among many atoms but generally the strongest at only one end of many of the same bonds as the VB state. Opposite symmetries enable optical transitions.

Spectral weights[38] from band unfolding further support the dominance of Γ states in both E^- and E^+ bands in Ge:C. The singly degenerate band at $E^- = E_{VBM} + 0.434$ eV shows a squared spectral weight of only 0.38 summed over all four L states and 0.34 at Γ. Similarly, the singly degenerate band at $E^+ = E_{VBM} + 0.825$ eV has squared spectral weights of 0.57 and 0.42 at Γ and L, respectively. These results still show that the Γ character of both E^+ and E^- states is comparable with or greater than their L character. These further suggest that mixing from L states does not dominate the properties of the C state. For reference, the sum of all squared spectral weights is nearly 1.00: 0.98 for E^- and 1.01 for E^+, showing that each state can be well represented by a basis set consisting of all Ge states that fold to Γ. Also, the relatively unperturbed, triply degenerate valence band maximum (without SOC) has a spectral weight of 2.9 at Γ.

The similarity between a Ge vacancy and a carbon atom in Ge becomes even more clear from Table III. Both types of defect shift the four Ge nearest neighbor atoms a comparable distance inward toward the defect, 12%–14% of the original 247 pm bond length. The angle from the second-nearest neighbors through the nearest neighbor to the defect is distorted from the original dihedral angle of 109.5° to roughly 103°, and the second-nearest bonds are conversely increased from 109.5° to just above 115°.

A significant contrast between Ge:C and v_{Ge} occurs in the top three VBs. A significant fraction of character of the top three Ge:C valence bands is shared with the v_{Ge} valence-like state deep within the bandgap. But this appears to be the v_{Ge} valence-like E_{a1} state picking up VB character rather than the other way around; the top Ge:C VBs consist almost entirely of unperturbed Ge VBs, as shown by the 100% overlap near $E = 0$. Because the v_{Ge} valence-like state

shares so much character with the Ge VB and increases the effective masses in the valence bands, there is an indication of BAC in the *valence* band of v_{Ge}, which we do not observe for Ge:C.

However, the influence of C does extend to the bottom of the VB. Projection of the very deepest valence band state ($\sim$11 eV below the bandgap) onto individual atoms shows that it is 42% on the carbon atom, despite C being only 0.78% of the alloy. This C contribution is 100.00% s-like in character. The addition of disorder (discussed below) further perturbs the VB states.

D. Strain and band anticrossing

It might be asked whether an artificially small bandgap in the computational model induces similar errors in the character of the E^- band; the Ge_{128} bandgap at Γ (0.788 eV) is smaller than the experimental one (0.80 eV), which could change the mixing of states in Ge:C. To address this question, we plotted the high-symmetry conduction bands (i.e., that fold to Γ) as a function of strain since strain shifts CB states at Γ more rapidly than at L or X. In Fig 5, marker positions show the bandgap for each CB state. Because Ge:C is highly perturbed and is not a simple repetition of the Ge primitive cell, the bands do not unfold to a clean, unique solution, and there is sometimes mixing between Γ, X, L, and other states. Each marker in Fig. 5 is itself a pie chart showing the proportions of L (red), Γ (green), and X (blue) symmetries from the unperturbed two-atom primitive cell contributing to that particular state. In other words, these measure symmetries of the wavefunction itself. This is different from Fig. 2, which shows overlaps with *Ge states* at different energies. Although the Ge conduction band minimum also has mostly spherical symmetry at each atom, it is not necessarily the same wavefunction as E^-, just as cos(r) and cos(2r) are both symmetric around r = 0 but they are not the same function.

The diameter of each marker in Fig. 5 represents the spectral weight of that state after unfolding. For example, states that bear little resemblance to the unperturbed two-atom primitive cell have small diameter markers, even if they have 100% overlap with other states. This allows a continuous identification of each band as the strain changes. Similarly, lines between markers have the same red–green–blue color dependence, and opacity $= \psi_{\varepsilon 1,k} | \psi_{\varepsilon 2,k} \cdot W_{\varepsilon 1} W_{\varepsilon 2}$, where $W_{\varepsilon 1}$ and $W_{\varepsilon 2}$ are the unfolded spectral weights of the end points at strains ε_1 and ε_2, respectively. In a perfect crystal, there would be only one opaque line connecting each marker, as shown for Ge128 in the inset. Despite the disorder, the lowest CB states from Γ (E-, lowest green), X (blue), and L (red) are all easily

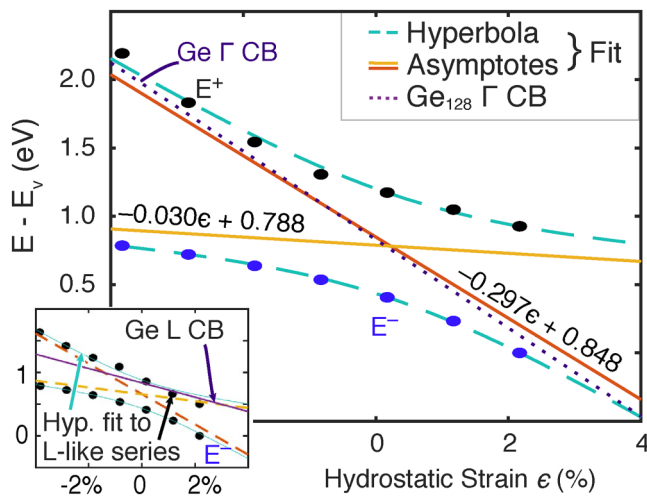**TABLE III.** Structural shifts due to C atom or Ge vacancy.

	Relaxed primitive lattice constant (nm)	Nearest atom shift toward defect (pm)	Distance between neighbors closest to defect (pm)	Angle through neighbor to defect (deg.)	Bond angle of 1st- to 2nd-nearest neighbors (deg.)
Ge (pure)	5.674	0	404	109.5 (all bonds)	109.5 (all bonds)
Ge vacancy	5.667	29	304	115.0	103.4
Ge:C	5.655	35	344	115.4	102.9
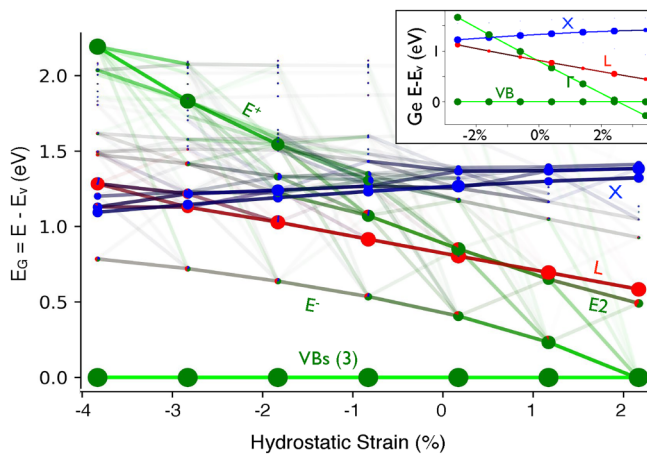
FIG. 5. Bandgap trends of several Ge:C conduction band states with increasing tensile strain. Individual markers are pie charts whose wedge areas and connecting line colors represent overlaps with L (red), Γ (green), and X (blue) states of the two-atom primitive cell. The valence bands ($E = 0$) are triply degenerate, so the markers are three times the normal area. They are also 100% Γ: green pie markers connected with green lines. The lowest conduction band (E^-) is mostly Γ as well. Inset: The same plot for Ge_{128}, which has no defect disorder.
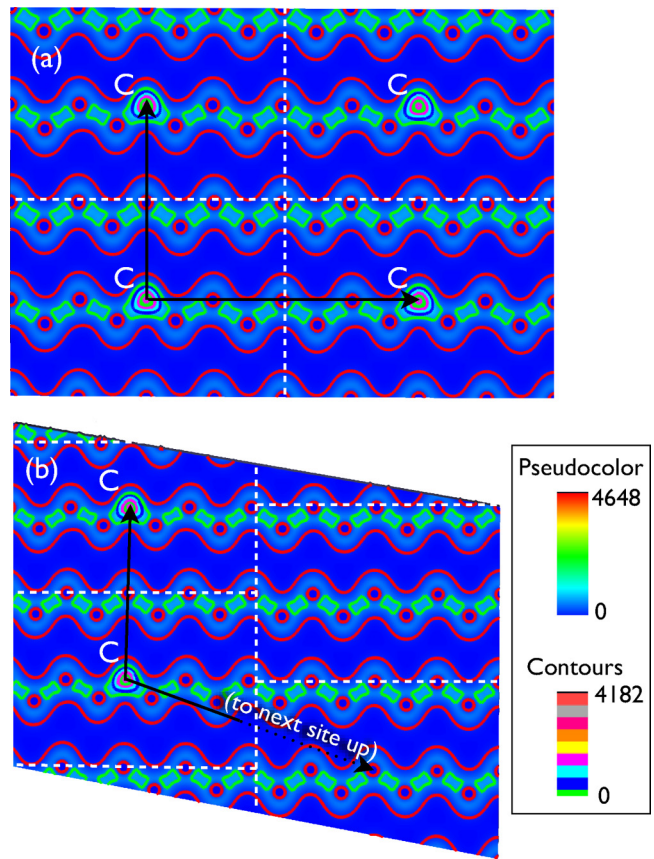
traceable with strain. All of these keep predominantly the same character over at least 0.5 eV of energy shift.

Figure 5 also shows two higher green bands, labeled E2 and E^+, with >95% Γ character that likewise decrease in energy with

strain. If the band anticrossing model applies, then one of these should anti-cross with the E^- band. Indeed, plotting only E+ with E− in Fig. 6 shows clear anticrossing behavior. A single hyperbola can fit both E+ and E− bands across the full range of strains, which strongly supports the band anticrossing model, at least for these high-symmetry points. The asymptotes to the hyperbola represent the original states before including the repulsive interaction between them. One state varies strongly with strain and is nearly identical with that of the unperturbed Ge CB. The other state is slowly varying, which we attribute to the isolated carbon "defect" level within the Ge CB, not to be confused with a midgap trap state. It is this slow variation with strain that Kirwan attributed to an L state, but which instead appears to be due to the C defect and mostly pinned to the vacuum level, almost independent of strain: just -30 meV per % strain, compared with 113 meV/%strain for the Ge L CBM.

FIG. 6. Ge:C E+ and E− bands from Fig. 5 (dots), along with fitted hyperbola (dashed lines) and its asymptotes (solid lines). The Ge Γ CB minimum is plotted as a dotted line for comparison and is almost identical to one asymptote. We interpret the other asymptote to be the isolated C "defect" state. Inset: Similar hyberbolic fits to the middle Γ-like series E2, which lies between E^+ and E^-, were unsuccessful; E2 is linear with strain.

FIG. 7. Charge density slice along (110) including multiple adjacent, cubic, 64-atom supercells with different basis vectors (black arrows). (a) x/y/z basis aligned with an eight-atom simple cubic cell, sliced through the C atom. (b) Staggered neighboring supercells, shifted by —one to two sites perpendicular to the plane. The C atom appears only in alternating atomic planes of the supercell.

As an alternative to E^+, noting that E^- becomes more L-like at strong compressive strains, we similarly attempted to fit the Γ-like (green) series E2 (between E^+ and E^-) with a hyperbola. However, the fit was very poor (Fig. 6, inset). The Ge L CB was also too high and the wrong slope to explain the strain dependence observed, which further suggests that the remaining L states do not mix or anti-cross with any other state, even though the Ge:C E^-. E^2. and E^+ states take on some L character. In contrast, the L-like series in Fig. 5 is quite linear, with a coefficient of determination $R^2 = 0.99$, and it tracks the Ge L CB almost exactly.

Based on these results, we identify the independent carbon "defect" in Ge:C to be approximately 0.788 eV above the VB at zero strain, decreasing by 30 meV per % hydrostatic strain. Like N in GaAs, the BAC effect is particularly strong because the C defect has almost exactly the same energy as the Γ valley in the unperturbed host, causing the strongest perturbation. It should be noted that the precise band energies may vary somewhat if empirical corrections are made to the model to match both L and Γ CB minima to experiment but such *ad hoc* parameter adjustment was not a focus of the present work.

E. Effects of supercell periodicity

Ge:C is expected to grow as a partly disordered alloy with C atoms no closer than the third-nearest neighboring sites from each other.[42] However, VASP's periodic boundary conditions impose a perfectly ordered supercell. To explore whether supercell periodicity was responsible for the charge anisotropy in Fig. 3, we modeled a 128 atom Ge:C system in which neighboring supercells were staggered with basis vectors (0.5, 0.75, 075), (0.625, 0.125, 0.5), and (0.5, 0.5, 0), so 1/3 of the supercell faces meet center to center, 1/3

meet face center to edge center, and 1/3 meet vertex to face center. As shown in Fig. 7, shifting the basis vectors had the effect of changing the distance and direction between C atoms. Although the anisotropy in charge density in Fig. 3 was not as apparent here, there was also no visible difference along the zigzag bonds in the {110} planes. In other words, the virtually identical Ge–Ge bond charge throughout the solid might not be due to a lack of interaction between neighboring C atoms, but rather due to electrons being shared *too well*, leading to a more uniform charge distribution throughout the supercell.

The conventionally stacked, face-to-face 128 atom supercells have a uniform distance from any C atom to its 12 nearest C neighbors in the neighboring supercells of 16.00 Å. The six second-nearest C neighbors are 22.62 Å apart. However, staggered supercells have more disorder: the nearest C–C neighbor distances are 14.45 Å, 15.52 Å, 16.03 Å, and 18.369 Å, depending on direction. This is reflected in the splitting of the lowest conduction bands into states with energies of 0.559, 0.805, 0.827, 1.241, 1.306, 1.329, and 1.369 eV above the valence band. The simulation of a true random alloy would require supercells much larger than are feasible at present.

This leaves the question why BAC fails to represent the band structure away from $k = 0$. For this, the E^- state at Γ was projected onto the full set of atoms [Figs. 8(a) and 8(b)], and the results are summed by the x–y position [Figs. 8(c) and 8(d)]. For conventional supercells with minimum length basis vectors, a weakly aligned charge density along ⟨111⟩ appears. This is also consistent with the 4th-nearest neighbor effects seen in Fig. 3. Using staggered supercells, where the center of one face meets a vertex or edge of the next, significant anisotropy appears along ⟨111⟩ directions.
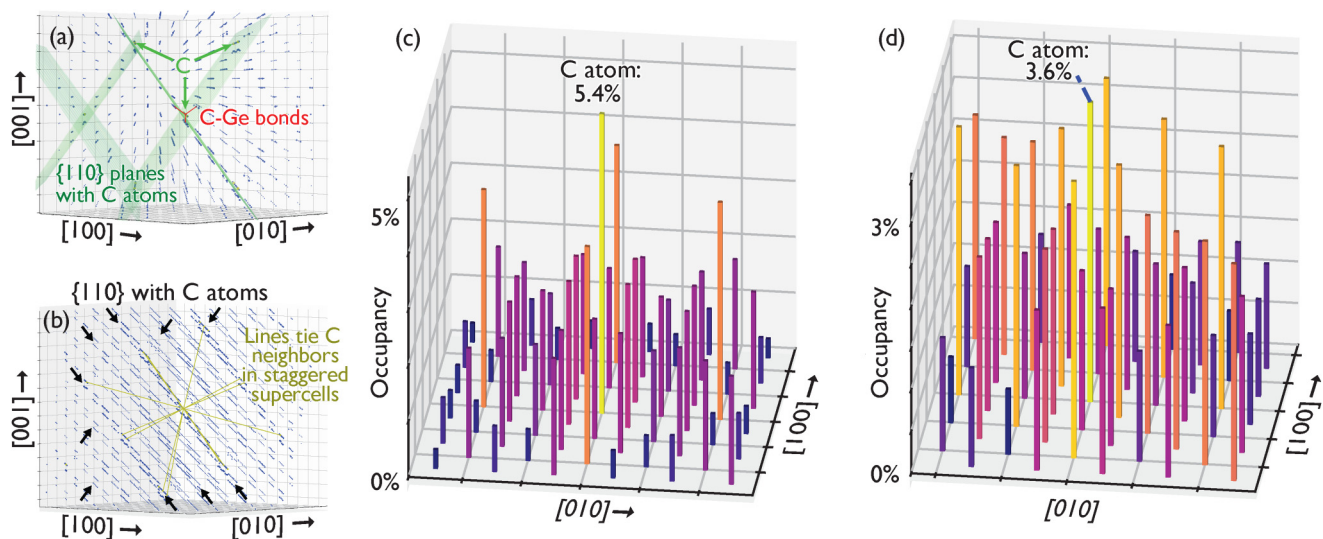
FIG. 8. Projection of the Ge:C E^- state at Γ (CBM) onto ions near the C atom for 128 atom supercells repeated with different basis sets. (a) and (b) Lattice view of ion occupation across multiple supercells. Several {110} planes containing C atoms are indicated by planes (a) or arrows (b). The spot diameter represents the fraction of wavefunction on atoms at that location. The dashed lines show C–Ge bonds. (c) and (d) Sum over vertically aligned atoms in (a) and (b), respectively, projected down onto the x–y plane ([100]–[010] plane). The bar height represents the fraction of state localized on atoms with the same x and y coordinates within the cube. (a) and (c) are from conventional supercell stacking (face to face); and (b) and (d) are from staggered supercells similar to Fig. 7(b).

TABLE IV. Projection of the lowest Γ conduction band state onto s, p, and d orbitals, summed over all atoms, shows negligible contribution from d electrons. The average VB–CB optical transition strengths (P_{cv}) at Γ are also comparable with and without SOC. The Ge_d:C SOC calculation included d electrons, spin–orbit coupling, and a smaller maximum radius for nonlocal projection operators (RMAX = 1.943 Å). Simulation conditions for Ge and Ge:C are the same as Fig. 1 with RMAX = 2.342 Å.

Material	s	p_y	p_z	p_x	d_{xy}	d_{yz}	d_{z2}	d_{xz}	d_{x2-y2}	Total	P_{cv}
Ge	0.137	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.137	9.8
Ge:C	0.515	0.030	0.030	0.030	0.011	0.011	0.001	0.011	0.001	0.641	4.9
Ge_d:C SOC	0.426	0.021	0.021	0.021	0.006	0.006	0.001	0.006	0.001	0.509	4.8

F. Anisotropy from periodic ordering

Even though the carbon "defect" is localized, its nonlocal charge in the E^- state, or at least the fraction projected onto atomic orbitals, extends four atoms away initially along $\langle 111 \rangle$ directions. This may account for the fractional L-like nature of the E^- state in Fig. 5 and reported by Kelires.[42] However, a closer inspection reveals that the charge is enhanced along {110} planes that include C atoms [Fig. 8(a)]. This is most prominent for the stacked super-cells, where C atoms occur only in every 4th bilayer. But stripes of excess charge are visible even in the staggered supercells, where C atoms occur in every other bilayer [Fig. 8(b)]. The anisotropy shown in Fig. 8 is not unique to the Ge-C system but it is some-what more delocalized than other HMAs such as GaAs:N or GaP:N.[43,44] We attribute the difference to the reduced difference in electronegativity: $\chi_B - \chi_{Ge} = 0.7$, compared with $\chi_N - \chi_{As} = 1.0$.

We interpret this preferential charge along $\langle 111 \rangle$ to be responsible for the divergence of band structures from the band anticrossing model when looking in different crystal directions. The difference in distribution of the E^- band over different atoms changes distinctly with different supercell basis vectors, even though the C atoms are always at least seven bonds away from each other. In light of this observation, it may be worth asking whether the anisotropic wavefunctions reported for similar HMAs are likewise due to the choice of supercell basis vectors. However, it is also worth noting that Fig. 8 does not plot the actual wavefunction probability density but the projection onto atomic orbitals. It may be possible that the orbitals along $\langle 111 \rangle$ simply line up with the E^- wavefunction better than others. Future work will study whether how such variations of a longer-ranged ordered alloy affect thermodynamic favorability, supercell shape, and band structure.

G. SOC and d electrons

Optical transitions from p-like to d-like states are also possible, so it is worth examining whether the strong optical transitions to the E^- band might instead be caused by a mixture of d-like states. Although d core electrons were not included in the PAW atomic *potentials* used above, the projection of final *wavefunctions* onto d orbitals would still indicate whether a tendency toward d orbitals was significant. However, a projection of the lowest conduction band onto Ge and Ge:C atomic orbitals shows very weak contribution from d orbitals, <2%, as shown in Table IV. The small fractions of the p and d characters in the E^- band do not explain the strong optical transitions remaining after adding C. We, therefore, believe Ge:C to be a true direct bandgap with largely s-like symmetry and strong optical transitions across the bandgap to and from

the two lowest conduction bands. Even if the fraction of C in the alloy was small enough that $E_{CB,L} < E_{CB,\Gamma}$, it would still be nearly direct rather than pseudo-direct.

One additional job was run including both d electrons and spin–orbit coupling (SOC). The precision of this calculation had to be reduced, with ENCUT = 410 eV and 374 conduction bands, to allow the job to converge within a reasonable time (several CPU decades). However, the orbital contributions and optical transition strengths were comparable with and without SOC, as shown in Table IV.

The use of pseudopotentials near atomic cores raises the question of the accuracy of projections of one state wavefunction onto another, and similarly the optical matrix elements for transitions between states. This is because pseudopotentials replace the actual potential near the atomic core with a constant or slowly varying potential in order to reduce computational demands. Because potentials are the deepest near the atom cores, the fraction of wavefunction near the cores may be non-negligible and oscillate strongly with position. This would reduce the accuracy of overlap integrals for the optical matrix elements. However, PAW potentials do reconstruct the exact valence wavefunction nodes near the atom cores.[28,29] This significantly increases the accuracy not only of band structures but also of the optical matrix elements, i.e., the overlap integral between the wavefunction of the initial state and the derivative of the wavefunction of the final state. Finally, we note that adding a vacancy strongly affects the valence bands, and it adds a VB-like state within the bandgap. This study was unable to include SOC or d electrons, which do affect the valence band, in v_{Ge}. Therefore, comparisons of Ge:C with v_{Ge} should be considered qualitative rather than quantitative.

IV. SUMMARY

In conclusion, 128 atom supercells of Ge:C with 1 C atom were modeled using hybrid functionals in order to study the origins of the new conduction bands, E^- and E^+, that are introduced by carbon. In contrast to recent reports, we find E^- and E^+ bands similar not only to each other, leading to band anticrossing, but also to the Ge Γ valley, and Ge:C shows optical transitions comparable with the Ge direct bandgap. L-valley conduction band states in Ge are ruled out as the major components of the E^- state in Ge:C by both a lack of change in the optical matrix elements across the bandgap at Γ and a smaller projection of E^- states onto L-valley states. Furthermore, spectral weights after band unfolding show comparable or more weight at Γ than at L for both E^+ and E^- states. These results were qualitatively similar whether we used

harder Ge potentials, compressive strain, or added spin–orbit coupling.

The pressure dependence of E^- nearly matches that of the L CB valley in Ge but this is found to be largely coincidental; both E^+ and E^- states come partly from orbitals farther from the bandgap that vary slowly in energy with pressure due to the carbon "defect" state being pinned to the vacuum level rather than the conduction band edge. This partly compensates the larger pressure dependence of Ge Γ CB. This is shown to be similar to vacancies in the Ge host matrix, in accordance with Hjalmarson's deep state model.

However, E^- is not well described by the band anticrossing model away from $k = 0$ nor does BAC explain a third CB state between E^+ and E^-. The differences from the first-order BAC model may be due to the bond distortion imposed by the introduction of C into the lattice or by the strong and anisotropic charge in the E^- band observed along $\langle 111 \rangle$ for some atom configurations, even with C atoms eight or more bonds away from each other. Although this could not be distinguished from super-periodicity imposed by the use of finite supercells in the present study, staggered supercells showed the same qualitative charge distribution and delocalization. It is noteworthy that both the physical range of filled-state charge displacement and the E^- CB state wavefunction reach at least as far as 4th-nearest neighboring atoms, in contrast to the strong localization reported by Kirwan et al.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] I. A. Gulyas and M. A. Wistey, "Optimum threshold current in double-gate tunnel FET devices" (unpublished).

[2] D. Gall, J. D'Arcy-Gall, and J. E. Greene, Phys. Rev. B 62(12), R7723 (2000).

[3] R. Soref, Silicon Photonics 8629, 862902 (2013).

[4] M. Wistey, Y.-Y. Fang, J. Tolle, A. V. G. Chizmeshya, and J. Kouvetakis, Appl. Phys. Lett. 90(8), 082108 (2007).

[5] C. A. Stephenson, W. A. O'Brien, M. W. Penninger, W. F. Schneider, M. Gillett-Kunnath, J. Zajicek, K. M. Yu, R. Kudrawiec, R. A. Stillwell, and M. A. Wistey, J. Appl. Phys. 120(5), 053102 (2016).

[6] J. Kouvetakis, A. Haaland, D. J. Shorokhov, H. V. Volden, G. V. Girichev, V. I. Sokolov, and P. Matsunaga, J. Am. Chem. Soc. 120(27), 6738–6744 (1998).

[7] J. D'Arcy-Gall, D. Gall, I. Petrov, P. Desjardins, and J. E. Greene, J. Appl. Phys. 90(8), 3910–3918 (2001).

[8] J. D'Arcy-Gall, D. Gall, P. Desjardins, I. Petrov, and J. E. Greene, Phys. Rev. B 62(16), 11203 (2000).

[9] S. B. Zhang and S.-H. Wei, Phys. Rev. Lett. 86(9), 1789 (2001).

[10] R. H. El-Jaroudi, K. M. McNicholas, B. A. Bouslog, I. E. Olivares, R. C. White, J. A. McArthur, and S. R. Bank, paper presented at the Conference on Lasers and Electro-Optics, San Jose, CA, 2019.

[11] S. R. Bank, H. P. Bae, H. B. Yuen, M. A. Wistey, L. L. Goddard, and J. S. Harris, Electron. Lett. 42(3), 156–157 (2006).

[12] M. A. Wistey, S. R. Bank, H. P. Bae, H. B. Yuen, E. R. Pickett, L. L. Goddard, and J. S. Harris, Electron. Lett. 42(5), 282–283 (2006).

[13] D. B. Jackrel, S. R. Bank, H. B. Yuen, M. A. Wistey, J. S. Harris, Jr, A. J. Ptak, S. W. Johnston, D. J. Friedman, and S. R. Kurtz, J. Appl. Phys. 101(11), 114916 (2007).

[14] E. D. Jones, N. A. Modine, A. A. Allerman, I. J. Fritz, S. R. Kurtz, A. F. Wright, S. T. Tozer, and X. Wei, Proc. SPIE 3621, 52–63 (1999).

[15] I. Gulyas, R. Kudrawiec, and M. A. Wistey, J. Appl. Phys. 126(9), 095703 (2019).

[16] S.Y. Park, J. D'Arcy-Gall, D. Gall, Y.-W. Kim, P. Desjardins, and J. E. Greene, J. Appl. Phys. 91(6), 3644–3652 (2002).

[17] S. Ismail-Beigi and S. G. Louie, Phys. Rev. Lett. 95(15), 156401 (2005).

[18] H. Koyama and K. Sueoka, J. Cryst. Growth 463, 110–115 (2017).

[19] H. P. Hjalmarson, P. Vogl, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44(12), 810 (1980).

[20] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids 44(5), 365–378 (1983).

[21] A. C. Kirwan, S. Schulz, and E. P. O'Reilly, Semicond. Sci. Technol. 34(7), 075007 (2019).

[22] J. L. Shay and J. W. Wernick, Ternary Chalcopyrite Semiconductors, Growth, Electronic Properties and Applications (Pergamon Press, London, 1975).

[23] I. A. Gulyas, C. A. Stephenson, and M. A. Wistey, "Convergence sensitivity of ab-initio models of a highly mismatched alloy, Ge:C" (unpublished).

[24] G. Kresse and J. Hafner, Phys. Rev. B 47(1), 558 (1993).

[25] G. Kresse and J. Hafner, Phys. Rev. B 49(20), 14251 (1994).

[26] G. Kresse and J. Furthmüller, Comput. Mater. Sci. 6(1), 15–50 (1996).

[27] G. Kresse and J. Furthmüller, Phys. Rev. B 54(16), 11169 (1996).

[28] P. E. Blöchl, Phys. Rev. B 50(24), 17953 (1994).

[29] G. Kresse and D. Joubert, Phys. Rev. B 59(3), 1758 (1999).

[30] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 77(18), 3865 (1996).

[31] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. 78(7), 1396–1396 (1997).

[32] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, J. Chem. Phys. 125(22), 224106 (2006).

[33] J. Hong, A. Stroppa, J. Íniguez, S. Picozzi, and D. Vanderbilt, Phys. Rev. B 85(5), 054417 (2012).

[34] S. H. Wei and A. Zunger, J. Vac. Sci. Technol. B 5(4), 1239–1245 (1987).

[35] L.-D. Yuan, H.-X. Deng, S.-S. Li, S.-H. Wei, and J.-W. Luo, Phys. Rev. B 98(24), 245203 (2018).

[36] R. M. Feenstra, N. Srivastava, Q. Gao, M. Widom, B. Diaconescu, T. Ohta, G. L. Kellogg, J. T. Robinson, and I. V. Vlassiouk, Phys. Rev. B 87(4), 041406 (2013).

[37] P. V. C. Medeiros, S. Stafström, and J. Björk, Phys. Rev. B 89(4), 041407 (2014).

[38] T. B. Boykin and G. Klimeck, Phys. Rev. B 71(11), 115215 (2005).

[39] T. B. Boykin, N. Kharche, G. Klimeck, and M. Korkusinski, J. Phys. Condens. Matter 19(3), 036203 (2007).

[40] M. Tomić, H. O. Jeschke, and R. Valentí, Phys. Rev. B 90(19), 195121 (2014).

[41] K. Momma and F. Izumi, J. Appl. Crystallogr. 44, 1272–1276 (2011).

[42] P. C. Kelires, Phys. Rev. B 60(15), 10837–10844 (1999).

[43] E. P. O'Reilly, A. Lindsay, and S. Fahy, J. Phys. Condens. Matter 16(31), S3257–S3276 (2004).

[44] P. R. C. Kent and A. Zunger, Phys. Rev. B 64, 115208 (2001).