RAPID COMMUNICATIONS

PHYSICAL REVIEW B VOLUME 36, NUMBER 8 15 SEPTEMBER 1987-I

# New optical transitions in strained Si-Ge superlattices

Sverre Froyen, D. M. Wood, and Alex Zunger
*Solar Energy Research Institute, Golden, Colorado 80401*
(Received 12 January 1987)

First-principles total-energy and electronic-structure calculations for $Si_nGe_n$ superlattices grown epitaxially on an (001) Si substrate reveal a nearly direct band gap despite the pronounced indirectness of its constituents. Whereas the lowest conduction-band wave function extends over both Si and Ge sublattices, the (higher energy) lowest direct state shows strong confinement to Si states. We predict that a substrate with a larger lattice constant than Si will produce a nearly direct band-gap superlattice (indirect only by 0.01 eV).

Many applications of "electronic materials" pose strict requirements on material properties, yet the handful of basic semiconductors provides only a discrete set of pertinent electronic properties. Whereas traditional chemical and biological technologies have created the needed diversity through direct synthesis of new species, semiconductor physics has largely taken the alternative route of material manipulations through alloying, creation of epitaxial strain, or growth of superstructures, using only a limited set of semiconducting "building blocks." One example is the recent observation of new direct transitions in strained $Si_nGe_n$ superlattices ($n=1-6$) composed of manifestly indirect band-gap Si and Ge units.[1] Another is the successful lowering of the conduction band of Si relative to that of the $Si_{0.5}Ge_{0.5}$ alloy in a strained $Si_n(Si_{0.5}Ge_{0.5})_m$ superlattice grown on (001) $Si_{0.75}Ge_{0.25}$, resulting in electron confinement in the Si layer and enhanced electron mobility.[2]

Using first-principles total-energy and electronic-structure calculations, we demonstrate how the combination of coherent strain with superlattice ordering can be used to manipulate the electronic structure of Si and Ge grown pseudomorphically on a (001) Si substrate.[1] We show the following. (i) Despite the much smaller gap of Ge relative to Si, the lowest direct conduction-band state of the strained $Si_nGe_n$ superlattice is confined to Si, not Ge. (ii) While this direct state is *localized* on Si, a *lower-energy* conduction-band state is *extended* on Si and Ge. (iii) Whereas the strained $Si_{0.5}Ge_{0.5}$ *alloy* has its lowest direct transition at[3] $\sim 2.6$ eV, the strained $Si_6Ge_6$ *superlattice* (differing from the alloy only in atomic ordering) has dramatically lower (by 1.6 eV) direct transitions. (iv) On a Si substrate the $Si_nGe_n$ superlattice still has an overall *indirect* band gap, but on a $Si_{0.5}Ge_{0.5}$ substrate the lowest direct gap is predicted to move to within 0.01 eV of the minimum indirect gap, making the system "quasi-direct."

We use the self-consistent pseudopotential total-energy method[4] within the local-density approximation (LDA) to calculate the equilibrium structure and band energies of (001) strained layer superlattices $Si_nGe_n$ for $n=1$, 2, 4, and 6. Nonlocal, semirelativistic pseudopotentials are generated by the method of Ref. 5. Wave functions are expanded in plane waves with kinetic energies up to 12 Ry for the calculation of total energies and 15 Ry for band energies. The calculation is performed self-consistently by evaluating the charge density at the equivalent of six special $\mathbf{k}$ points (ten for the $n=4$ superlattice) in the irreducible face-centered-cubic (fcc) Brillouin zone. Exchange and correlation are treated using the electron gas data of Ceperley and Alder, as parametrized by Perdew and Zunger.[6] Equilibrium geometries are determined through first-principles total-energy minimizations.[4] We estimate that band energy errors caused by incomplete structural relaxation are less than 0.03 eV. Convergence errors are of similar magnitude. The LDA leads to the usual underestimates of conduction-band energies, but, as shown below, our conclusions will depend only on the behavior of the $X$ point (and to a lesser degree the $L$ point) in the fcc Brillouin zone. Since LDA errors at $X$ and $L$ for both Si and Ge are similar (0.6 eV), we expect that this error will not alter our conclusions.

We describe strain- and ordering-induced changes in the electronic structure in three conceptual steps: (i) effects of pressure and strain on the pure constituents, (ii) the effect of alloying these constituents, and (iii) ordering of the alloy into a superlattice. In what follows we will label superlattice states evolving from the fcc state $\Gamma$, $X$, and $L$ as $\bar{\Gamma}$, $\bar{X}$, and $\bar{L}$, respectively.

*Effect of pressure and strain on the pure constituents.* Pure Ge layers can be grown in a pseudomorphic, dislocation-free manner on a Si substrate provided the strain resulting from the lattice mismatch can be accommodated elastically, limiting the thickness to a critical value,[7] typically a few monolayers. This results in compression of the cubic Ge lattice constant $a_{Ge}$ to that ($a_{Si}$) of Si (a hydrostatic pressure effect) and a tetragonal elongation of the $c$ axis ($c/a>1$). We note first that for cubic Ge ($c/a=1$) at $a_{Si}$ [Fig. 1(b)] the hydrostatic part of the strain shifts the $L_c$ and $\Gamma_c$ conduction ($c$) band states of Ge (having positive deformation potentials[8]) to *higher* energies relative to the valence ($v$) band maximum [Figs. 1(a) and 1(b)]. In contrast, since the $X_c$ state has a negative deformation potential,[8] its energy is *lowered*. We find that cubic Ge at $a_{Si}$ has an indirect gap via $X_c$, and not $L_c$ [Fig. 1(b)]. Second, turning to tetragonal deformation effects, total-energy minimizations for Ge and $Si_nGe_n$ lattice-matched to (001) show a 6% increase in $c/a$ for Ge, and a 3% increase for $Si_nGe_n$ for $n=1$, 2, 4, and 6. This biaxial compressive strain (along the [001]

 

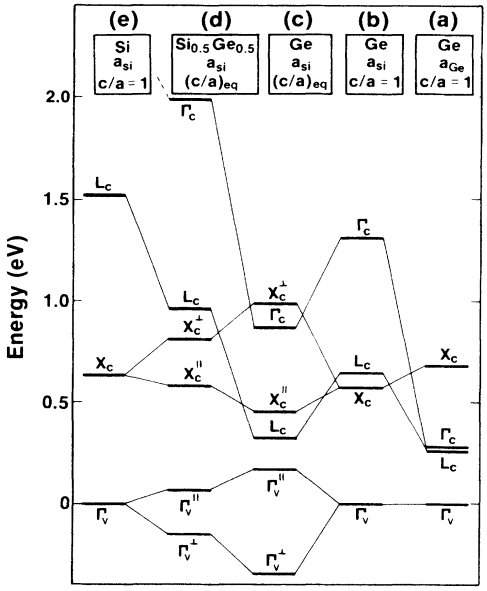


FIG. 1. (a) Energy levels (in eV, relative to the average of the strain-split valence-band maximum states) at $\Gamma$, $X$, and $L$ of (a) bulk Ge ($a_{Ge}=5.61$ Å) and (e) Si ($a_{Si}=5.41$ Å), as well as (d) distorted Ge [(b) and (c)] and a $Si_{0.5}Ge_{0.5}$ alloy ($a_{alloy}=5.51$ Å).

and [010] directions) splits the six $X_c$ conduction-band valleys of Ge into two, degenerate longitudinal states $X_c^{\perp}$ and four, degenerate transverse states $X_c^{\parallel}$ [Fig. 1(c)]. For $(c/a)>1$ (appropriate here) the $X_c^{\perp}$ states shift to higher energy relative to $X_c^{\parallel}$ (a splitting of 0.5 eV for Ge). In contrast, since the Ge conduction-band valleys at $L_c$ are oriented along the [111] directions, they remain unsplit by the strain and shift up uniformly, resulting in near degeneracy of $L_c$ and $X_c^{\parallel}$ [Fig. 1(c)]. The threefold-degenerate valence-band maximum splits uniaxially into an upper, doubly degenerate state and a lower, singly degenerate state. Tetragonal deformation, therefore, results in a *reduction* of the indirect band gap and an *increase* of the direct band gap of epitaxial Ge relative to its bulk form [Fig. 1(c) versus Fig. 1(a)]. The objective of making the system more direct is, hence, not met by epitaxial effects alone.

*Alloying effects.* Further manipulation of the electronic structure of the Si-Ge system has been sought through epitaxial alloying.[2,3] We model this system by a virtual crystal calculation for $Si_{0.5}Ge_{0.5}$ matched epitaxially to $a_{Ge}$. The resulting energy levels [Fig. 1(d)] are close to the average of Si [Fig. 1(e)] and compressed, tetragonally distorted Ge [Fig. 1(c)]. Like Si, the strained 50%-50% alloy has its conduction-band minimum at $X$ [Fig. 1(d)]. However, the $\Gamma_v^{\parallel}\rightarrow X_c^{\parallel}$ minimum gap of the *strained* alloy is lower (by $\sim$0.1 eV) than that of the *unstrained* alloy [approximately the average of Figs. 1(a) and 1(e)]. The lowest direct $\Gamma_v\rightarrow\Gamma_c$ transition of the alloy occurs at a considerably higher energy than the indirect gap (LDA result: 1.9 eV; experimental:[3] 2.6 eV). Hence, the strained alloy is only marginally more direct than Si. We may quantify these effects using the band-gap indirectness parameter $\delta\equiv 2(E_g^D-E_g^I)/(E_g^D+E_g^I)$ for the direct $\Gamma_v\rightarrow\Gamma_c$ band gap ($E_g^D$) relative to the indirect $\Gamma_v\rightarrow X_c$ band gap ($E_g^I$). We find $\delta=1.16$ for the epitaxial alloy, comparable to $\delta=1.20$ for Si.

*Ordering the alloy into a superlattice.* Finally, we contrast the properties of the disorder *alloy* with those of the artificially ordered[1] $Si_nGe_n$ *superlattice*. To compute the electronic structure of strained $Si_nGe_n$, one must first find equilibrium atomic positions. For a fixed $a_{Si}$ (calculated to be 5.41 Å), these were found by minimizing the total energy through simultaneous optimization of $c/a$ and the cell-internal atomic positions.[9] We find the spacings between adjacent planes $\alpha$ and $\beta$, $d_{\alpha\beta}=(c/2n)(1+\epsilon)$ ($\epsilon=0$ for the unrelaxed structure), relax so that $d_{Si\text{-}Si}$ is close to its value in bulk Si ($a_{Si}/4$, i.e., $\epsilon\simeq -0.03$), while $d_{Ge\text{-}Ge}$ is close to its value in epitaxial Ge on Si ($1.06a_{Si}/4$, i.e., $\epsilon\simeq +0.03$). Relaxation affects both the electronic structure (much as bond alternation in alloys affects their band gaps[10]) and the system's total energy. The superlattice formation enthalpy is found to be $0\pm 0.05$ kcal/atom pair in the relaxed *epitaxial* systems (all $n$'s), while the formation enthalpy relative to the *bulk* end-point compounds is considerably larger, +0.7 kcal/pair.

Given this enhanced stability of the epitaxial superlattice, we explore the consequences of artificial atomic ordering[1] and symmetry lowering for its electronic structure. Ordering not only further reduces the minimum band gap, but also dramatically renders it "quasidirect" (i.e., $\delta\simeq 0.1$ for the $n=6$ superlattice on a Si substrate). The increased cell dimension in the [001] direction perpendicular to the layers leads to the usual folding[11] of states along the $\Gamma$-$X$ line of the fcc zone into the center $\bar{\Gamma}$ of the superlattice zone. If the $X_c$ states of the constituents are lower in energy than their $\Gamma_c$ states [Figs. 1(c) and 1(e)] this leads to a large *increase* in the degree of directness of the band gap. Indeed, lattice-matched $(AlAs)_n(GaAs)_n$ superlattices[12] ($c/a\simeq 1$) are direct for $n>1$ despite the fact that the 50%-50% alloy is indirect. However, strain-induced tetragonal distortion permits folding into $\bar{\Gamma}$ of the longitudinal $X_c^{\perp}$ states but not of the transverse $X_c^{\parallel}$ state. Since for $Si_nGe_n$ on Si the tetragonal distortion ($c/a>1$) shifts the folding $X_c^{\perp}$ state *above* the nonfolding $X_c^{\parallel}$ state [Fig. 1(c)], it reduces the directness. Thus there are two competing effects on band-gap directness: folding, which *increases* directness, and tetragonal distortion, which will make the system *less* direct. A detailed self-consistent calculation is required to determine the balance between these trends. Such a calculation exhibits *selective quantum confinement effects* associated with the different potentials in the Si and Ge regions. In particular, we find that energy levels of the superlattices naturally divide into two categories [Figs. 2(a) and 2(b), respectively].

The first group of states [Fig. 2(a)] is superlattice states at $\bar{\Gamma}$ which evolve from folding of the $X_c^{\perp}$ states. These behave as conventional quantum-confined "Kronig-Penney" states[11] in two respects. First, their energies drop as the well width (the repeat period $n$) increases. Second, the wave function amplitudes for the lowest members of the series [$\bar{\Gamma}_1(X_c^{\perp})$ and $\bar{\Gamma}_2(X_c^{\perp})$ in Figs. 3(f) and 3(g)] are confined almost exclusively to the Si well, reflecting strain-induced lowering[13] of the conduction-

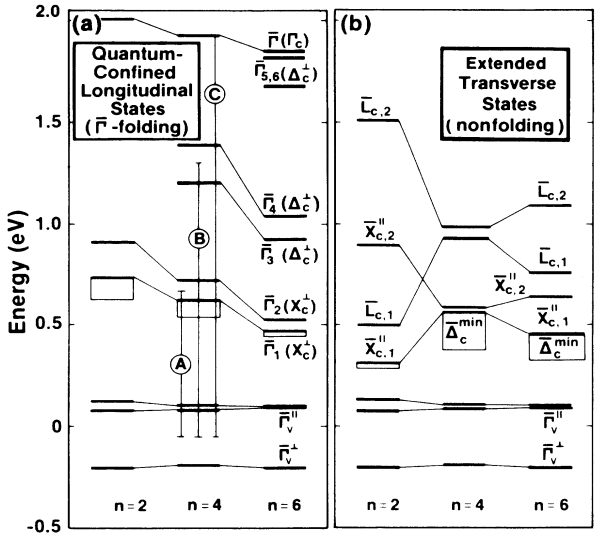


FIG. 2. (a) Energy levels (in eV) of quantum-confined longitudinal states and (b) extended transverse states of strained (001) $Si_nGe_n$ superlattices grown on Si for $n=2$, 4, and 6. Boxed regions indicate the extent of downward dispersion of a band away from the symmetry point. The zero of energy is taken as the average of the top three valence states at $\bar{\Gamma}$.

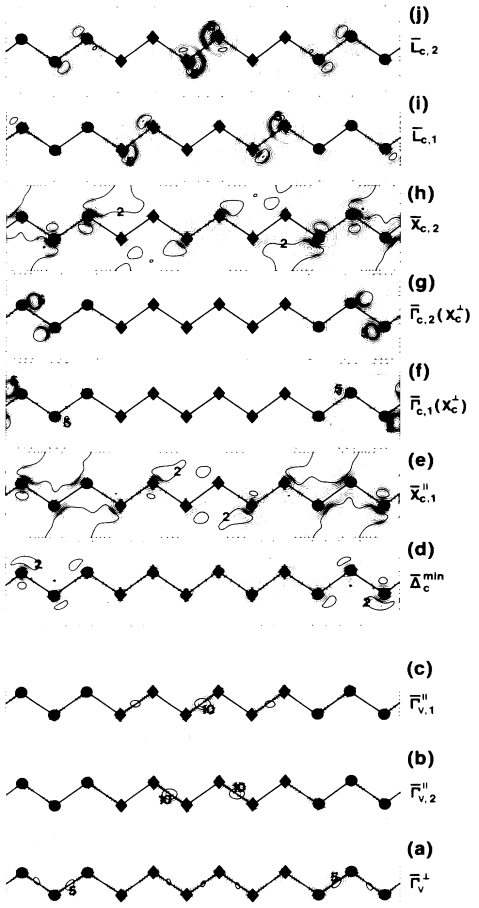


FIG. 3. Contour plots of the probability density for states at $\bar{\Gamma}$, $\bar{X}$, and $\bar{L}$ (in increasing order of energy) in the $Si_6Ge_6$ superlattice. Atomic positions are marked with solid diamonds (Ge) and circles (Si).

band minimum of Si relative to that of Ge and the *heavy* longitudinal ($l$) electron mass for Si, $m_e^l(\mathrm{Si})=0.98$.

These folding longitudinal states result in new direct transitions with no counterpart in the random alloy. They are denoted as $A$ and $B$ in Fig. 2(a) (for $n=4$) and have calculated onsets at 0.6 ($A$) and 1.1 eV ($B$). The transition labeled $C$ [calculated: 2.0 eV; observed (Ref. 1): 2.31 eV] is the superlattice analog of the direct alloy $\Gamma_v \rightarrow \Gamma_{2'c}$ transition [Fig. 1(d)]. While $C$ is thus direct, $A$ and $B$ are "pseudodirect," i.e., arise from a folded state. Atomic relaxation and charge redistribution are found to make transitions $A$ and $B$ of considerably larger intensity than truly indirect transitions. We find that the ratio $f_\alpha/f_C$ of the energy- and mass-weighted dipole matrix elements appropriate to electroreflectance is[14] $\sim 0.03$ for $\alpha=A$ and $\sim 0.1$ for $\alpha=B$, for all polarizations.

The second group of states [Fig. 2(b)] is the transverse states $\bar{X}_c^{\parallel}$ and $\bar{L}_c$ which do not fold into $\Gamma$.[15] These states [Figs. 3(e) and 3(h)–3(j)] are extended over *both* Si and Ge sublattices because of their much *lighter* transverse ($t$) electron masses [$m_e^t(\mathrm{Si},X_c)=0.19$ and $m_e(\mathrm{Ge},L_c)=0.12$].[16] Therefore, in $Si_nGe_n$ we encounter the unusual situation where the *lower* energy (transverse) states [Fig. 3(e)] are *extended* and *higher* energy (longitudinal) states [Figs. 3(f) and 3(g)] are *localized*. Since these transverse states are extended, their variation with $n$ for $Si_nGe_n$ [Fig. 2(b)] does not follow the monotonic decrease with $n$ characteristic of the quantum-confined states [Fig. 2(a)]; in fact, they are seen to *oscillate* with $n$. Each pair of states $\{\bar{X}_{c1}^{\parallel},\bar{X}_{c2}^{\parallel}\}$ (Ref. 17) and $\{\bar{L}_{c1},\bar{L}_{c2}\}$ in Fig. 2(b) derives from the corresponding $X_{1c}$ and $L_{1c}$ state, respectively, in Si and Ge. Since they are extended over both sublattices, the centroid for each pair is approximately independent of $n$. In contrast to the *alloy* [Fig. 1(d)], however, each pair in the *superlattice* is split by the different potentials in the Si and Ge regions. The variation of this splitting (i.e., the oscillations) with $n$ may be understood from inspection of their wave functions for, e.g., $n=6$ [Figs. 3(e) and 3(h)–3(j)].[18] These states show a quasi-periodicity of four atomic monolayers in the [001] direction, reflecting the underlying atomic repeat period for the diamond structure in this direction. For $n=2$, this period matches that of the superlattice, permitting the states to selectively localize on either the Si or Ge sublattice, yielding a *maximum* splitting. For $n=4$ the states must sample both sublattices, resulting in almost *no splitting* [Fig. 2(b)].

We turn finally to the factors determining the direct/indirect nature of the conduction band of lowest energy for $Si_nGe_n$. This is decided (in order of importance) by (i) the magnitude of the tetragonal distortion, (ii) the strong $n$ dependence of the longitudinal and transverse states, and (iii) possible downward dispersion away from high-symmetry points in the zone (as in Si around the $X$ point). We note that (iii) both $\bar{\Gamma}$ and $\bar{X}$ derive from $X$

and, indeed, show such an effect; the magnitude of the dispersion is shown by boxed areas in Fig. 2. For a Si substrate we find indirect minima lower by 0.4, 0.2, and 0.1 eV than the lowest direct gaps for $n=2$, 4, and 6, respectively. Despite this indirectness, ordering dramatically enhances the degree of band gap directness; we find $\delta=1.08$, 0.11, and 0.03 for $n=2$, 4, and 6, respectively.

The preceding discussion suggests how the situation above, with a Si substrate, may be modified in order to make $Si_nGe_n$ more nearly direct. The principal reason [(i) above] for indirectness is that the relatively large value of $c/a$ displaces the folding $\bar{X}_c^{\perp}$ to higher energy than the nonfolding $\bar{X}_c^{\parallel}$ state [Figs. 1(c) and 1(d)]. Repeating our calculations for $n=4$ using $Si_{0.5}Ge_{0.5}$ as a substrate, we find $c/a \simeq 1$. The conduction-band minimum at $\bar{\Delta}_c^{\min}$ close to $\bar{X}_{c1}^{\parallel}$ is now 0.04 eV higher than $\bar{\Gamma}_1(X_c^{\perp})$ and the only remaining indirectness is limited to the small downward dispersion away from $\bar{\Gamma}$ along $k_{\perp}$ (0.1 eV for $n=4$, 0.01 eV for $n=6$). This may be further reduced by growing $Si_nGe_n$ for $n>6$, which in turn will be facilitated by the smaller lattice mismatch between this substrate and the constituents. Since the average mismatch is zero, it furthermore permits growth of superlattices of optical thickness. Experiments on $Si_nGe_n$ for a 50%-50% alloy substrate are called for.

The authors are grateful to T. P. Pearsall for providing us with his experimental results (Ref. 1) prior to publication.

[1]T. P. Pearsall, J. Bevk, L. C. Feldman, A. Ourmazd, J. M. Bonar, and J. P. Mannaerts, Phys. Rev. Lett. **58**, 729 (1987).

[2]G. Abstreiter, H. Brugger, T. Wolf, H. Jorke, and H. J. Herzog, Phys. Rev. Lett. **54**, 2441 (1985).

[3]T. P. Pearsall, F. H. Pollak, J. C. Bean, and R. Hull, Phys. Rev. B **33**, 6821 (1986).

[4]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4404 (1979).

[5]We use "norm-conserving" pseudopotentials of G. Kerker, J. Phys. C **13**, L189 (1980) [essentially identical in construction to those of A. Zunger and M. L. Cohen, Phys. Rev. B **20**, 4082 (1979), termed there analytically continued potentials].

[6]See J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981), and references therein.

[7]T. Narusawa and W. M. Gibson, Phys. Rev. Lett. **47**, 1459 (1981).

[8]F. H. Pollak and M. Cardona, Phys. Rev. **172**, 816 (1968).

[9]Previous model calculations on $Si_nGe_n$ superlattices, e.g., U. Gnutzmann and K. Clausecker, Appl. Phys. **3**, 9 (1974); and J. A. Moriarty and S. Krishnamurthy, J. Appl. Phys. **54**, 1892 (1983), have ignored $c/a$ distortions, atomic relaxations, and self-consistency.

[10]A. Zunger and J. E. Jaffe, Phys. Rev. Lett. **51**, 662 (1983).

[11]L. Esaki and R. Tsu, IBM J. Res. Dev. **14**, 61 (1970).

[12]A. Ishibashi, Y. Mori, M. Itabashi, and N. Watanabe, J. Appl. Phys. **58**, 2691 (1985).

[13]C. G. Van der Walle and R. M. Martin, J. Vac. Sci. Technol. B **4**, 1055 (1986).

[14]These values are sensitive to atomic relaxation.

[15]$Si_nGe_n$ for $n$ even has two inequivalent $L$ points. In the text we discuss the more important, which is split by the superlattice potential into $\bar{L}_{c1}$ and $\bar{L}_{c2}$, and whose $k$ vector is along Si—Ge bonds. The second $L$ point has an unsplit doubly degenerate conduction-band minimum with an energy equal to the average of $\bar{L}_{c1}$ and $\bar{L}_{c2}$. These states are also extended, with the same effective mass as $\bar{L}_{c1,2}$.

[16]The appropriate mass for the $L$ point is the effective mass in the [001] direction. Valence-band states [Figs. 3(a)–3(c)] are extended (with excess amplitude on Ge) since they are characterized by the intermediate Ge hole mass, $m_h(\mathrm{Ge})=0.3$.

[17]The states $\bar{X}_{c1}^{\parallel}$ and $\bar{X}_{c2}^{\parallel}$ in Fig. 2(b) are, in fact, each doubly degenerate; in the text we do not make this distinction.

[18]We find the energies of the $\bar{X}_v$ and $\bar{L}_v$ at $E_{\mathrm{VBM}}-3$ and $E_{\mathrm{VBM}}-0.9$ eV, respectively, rather $n$ dependent; the $\bar{L}_v \rightarrow \bar{L}_c$ and $\bar{X}_v \rightarrow \bar{X}_c^{\parallel}$ transitions, hence, provide a useful gauge of oscillations of conduction states with $n$.