

Supplementary Material for "Conduction band nonparabolicity, chemical potential, and carrier concentration of intrinsic InSb as a function of temperature"

(Dated: 21 November 2024)

Stefan Zollner,¹ Carlos A. Armenta,¹ Sonam Yadav,¹ and José Menéndez²

¹*Department of Physics, New Mexico State University, MSC 3D, P.O. Box 30001, Las Cruces, NM 88003, USA*

²*Department of Physics, Arizona State University, Tempe, AZ 85287-1504, USA*

S1. QUANTUM MECHANICAL PERTURBATION THEORY

A. Non-degenerate perturbation theory

Let us assume [following the Wikipedia article *Perturbation theory (quantum mechanics)*] that a quantum mechanical system is described, for the most part, by a Hamiltonian H_0 which has a finite number of eigenstates $|n0\rangle$ with eigenvalues E_{n0} given by

$$H_0 |n0\rangle = E_{n0} |n0\rangle; \quad n = 1, \dots, N. \quad (\text{S1})$$

In addition, there is a small perturbation $H_{\vec{k}}$ that depends on some vector parameter \vec{k} , which can be made arbitrarily small. The total Hamiltonian is therefore $H = H_0 + H_{\vec{k}}$ with eigenstates $|n\vec{k}\rangle$ and eigenvalues $E_{n\vec{k}}$ given by

$$H |n\vec{k}\rangle = E_{n\vec{k}} |n\vec{k}\rangle; \quad n = 1, \dots, N. \quad (\text{S2})$$

To calculate the perturbed eigenvalues and eigenstates of H for small values of $|\vec{k}|$, we must distinguish between two separate cases, depending on whether the eigenstate $|n0\rangle$ is non-degenerate or degenerate.

In the **non-degenerate** case, the perturbed eigenvalues and eigenstates are given to **first order** in the perturbation by

$$E_{n\vec{k}} = E_{n0} + \langle n0 | H_{\vec{k}} | n0 \rangle, \quad (\text{S3})$$

$$|n\vec{k}\rangle = |n0\rangle + \sum_{i \neq n} |i0\rangle \frac{\langle i0 | H_{\vec{k}} | n0 \rangle}{E_{n0} - E_{i0}}. \quad (\text{S4})$$

We sometimes also need to know the eigenvalues to second order in $|\vec{k}|$, which are given in **second order** perturbation theory by

$$E_{n\vec{k}} = E_{n0} + \langle n0 | H_{\vec{k}} | n0 \rangle + \sum_{i \neq n} \frac{|\langle i0 | H_{\vec{k}} | n0 \rangle|^2}{E_{n0} - E_{i0}}. \quad (\text{S5})$$

These equations can be written in more compact form, if we introduce the notation

$$H_{ni} = \langle n0 | H_{\vec{k}} | i0 \rangle \quad \text{and} \quad E_{ni} = E_{n0} - E_{i0}. \quad (\text{S6})$$

The first order perturbation theory results in

$$E_{n\vec{k}} = E_{n0} + H_{nn}, \quad (\text{S7})$$

$$|n\vec{k}\rangle = |n0\rangle + \sum_{i \neq n} |i0\rangle \frac{H_{in}}{E_{ni}} \quad (\text{S8})$$

and second order perturbation theory in

$$E_{n\vec{k}} = E_{n0} + H_{nn} + \sum_{i \neq n} \frac{|H_{ni}|^2}{E_{ni}}. \quad (\text{S9})$$

B. Degenerate perturbation theory

If the energy level E_{n0} is **degenerate**, then Eq. (S3) is ambiguous, because H_{nn} depends on the choice of the basis and therefore may not be uniquely defined. Also, the energy denominator E_{ni} will vanish if $|i0\rangle$ is in the degenerate eigenspace belonging to eigenvalue E_{n0} and therefore the second-order energy correction (S5) will become singular.

We address this by assuming that g is the degeneracy of the E_{n0} eigenspace of the unperturbed Hamiltonian. If we choose arbitrary basis functions $|i0\rangle$ for this eigenspace, then the wave function $|n0\rangle$ can be written as a linear combination

$$|n0\rangle = \sum_{i=1}^g c_i |i0\rangle \quad (\text{S10})$$

with some coefficients c_i . The perturbed energies to first order perturbation theory then become

$$E_{n\vec{k}} = E_{n0} + E_i, \quad (\text{S11})$$

where the E_i are the eigenvalues of the perturbation matrix $H_{ij} = \langle i0 | H_{\vec{k}} | j0 \rangle$ within the degenerate eigenspace of E_{n0} . The perturbation will usually lift some, but not necessarily all of the degeneracies of this eigenspace.

If the perturbation is sizeable and the corrected energies encroach on nearby energy levels of the unperturbed Hamiltonian, it may be necessary to also include the nearby levels in the calculation. In first-order perturbation theory, the perturbed energies are then given by the eigenvalues of the matrix

$$\langle i0 | H_0 + H_{\vec{k}} | j0 \rangle, \quad (\text{S12})$$

where as many eigenstates of H_0 as needed can be included in the matrix.

S2. $\vec{k} \cdot \vec{p}$ -THEORY

Electron states in solids are described by three quantum numbers, the wave vector \vec{k} , the band index n with energy $E_{n\vec{k}}$, and the spin s (which we ignore for the moment). The one-electron Hamiltonian H is the sum of the kinetic energy derived from the momentum operator $\vec{p} = -i\hbar\vec{\nabla}$ and the crystal potential $V(\vec{r})$. If $\phi_{n\vec{k}}$ is the wave function, then the Schrödinger equation is

$$H\phi_{n\vec{k}} = \left[\frac{\vec{p}^2}{2m_0} + V(\vec{r}) \right] \phi_{n\vec{k}} = E_{n\vec{k}}\phi_{n\vec{k}}, \quad (\text{S13})$$

where m_0 is the free electron mass. Bloch's theorem allows us to write the wave function as the product of a plane wave and a function $u_{n\vec{k}}$, which is periodic in the crystal lattice:

$$\phi_{n\vec{k}} = e^{i\vec{k}\cdot\vec{r}} u_{n\vec{k}}. \quad (\text{S14})$$

If we insert Eq. (S14) into the Schrödinger equation (S13) and use the product rule for the Laplace operator $\Delta f g = f\Delta g + 2\vec{\nabla}f \cdot \vec{\nabla}g + g\Delta f$, we obtain a new equation for the periodic part of the wave function, see Yu and Cardona (2.35):

$$\left[\frac{\vec{p}^2}{2m_0} + \frac{\hbar}{m_0}\vec{k} \cdot \vec{p} + \frac{\hbar^2 k^2}{2m_0} + V(\vec{r}) \right] u_{n\vec{k}} = E_{n\vec{k}} u_{n\vec{k}}. \quad (\text{S15})$$

For a vanishing wave vector $\vec{k}=0$, the kinetic energy and potential terms have solutions that are assumed to be known, for example from optical measurements of the gaps at critical points, see Yu & Cardona (2.36),

$$\left[\frac{\vec{p}^2}{2m_0} + V(\vec{r}) \right] u_{n0} = E_{n0} u_{n0}. \quad (\text{S16})$$

The \vec{k} -dependent terms in Eq. (S15) are treated in perturbation theory up to second order in \vec{k} .

S3. KANE'S SUBSTITUTION

The physical significance of the four terms in square brackets in Eq. (S15) is as follows: The first and fourth terms are the kinetic and potential energy of the electron at the Γ -point ($\vec{k}=0$) belonging to the unperturbed Hamiltonian. The second term is the $\vec{k} \cdot \vec{p}$ -term, which will be treated in perturbation theory. The third term, which is quadratic in k , is called the free electron term, because it describes the kinetic energy of a free electron with mass m_0 as a function of crystal momentum $\hbar\vec{k}$.

Kane (1956, 1957) notes that the formalism can be simplified by introducing a **modified energy parameter**

$$\tilde{E}_{n\vec{k}} = E_{n\vec{k}} - \frac{\hbar^2 k^2}{2m_0}, \quad (\text{S17})$$

where the free electron kinetic energy has been subtracted. We use a tilde (instead of a prime) to denote the modified energies, because the prime (as in E'_0) has a different meaning in the critical-point notation proposed by Cardona.

This simplifies Eq. (S15) to

$$\left[\frac{\vec{p}^2}{2m_0} + \frac{\hbar}{m_0}\vec{k} \cdot \vec{p} + V(\vec{r}) \right] u_{n\vec{k}} = \tilde{E}_{n\vec{k}} u_{n\vec{k}}. \quad (\text{S18})$$

The $\vec{k} \cdot \vec{p}$ -term is then treated as a perturbation.

For a non-degenerate band, the first-order correction is

$$\tilde{E}_{n\vec{k}} = E_{n\vec{k}} - \frac{\hbar^2 k^2}{2m_0} = E_{n0} + \frac{\hbar}{m_0} \langle n0 | \vec{k} \cdot \vec{p} | n0 \rangle. \quad (\text{S19})$$

The term with the dot product is linear in \vec{k} . This term must vanish in crystals with inversion symmetry. It also vanishes in crystals without inversion symmetry due to the Kramers degeneracy (time-reversal symmetry), if spin effects are neglected. Note: There are small \vec{k} -linear terms in the band structure of zinc blende semiconductors, but they arise from relativistic corrections, not from the last term in Eq. (S19).

The perturbed wave function for non-degenerate bands is derived from Eq. (S4) and found to be, compare Yu & Cardona (2.37),

$$u_{n\vec{k}} = u_{n0} + \frac{\hbar}{m_0} \sum_{i \neq n} |i0\rangle \frac{\langle i0 | \vec{k} \cdot \vec{p} | n0 \rangle}{E_{n0} - E_{i0}}. \quad (\text{S20})$$

The perturbed wave functions in Eq. (S20) are needed to calculate the correction to the energy in second order perturbation theory with Eq. (S5).

Since the first-order energy correction vanishes, we proceed with the correction in second-order perturbation theory and find

$$\tilde{E}_{n\vec{k}} = E_{n0} + \frac{\hbar^2}{m_0^2} \sum_{i \neq n} \frac{|\langle n0 | \vec{k} \cdot \vec{p} | i0 \rangle|^2}{E_{n0} - E_{i0}} \quad (\text{S21})$$

for non-degenerate levels E_{n0} .

Similarly, we need to diagonalize the matrix

$$\sum_i \left(E_{n0} \delta_{ni} + \frac{\hbar}{m_0} \langle n0 | \vec{k} \cdot \vec{p} | i0 \rangle \right) c_{ni} = \tilde{E}_{n\vec{k}} c_{n\vec{k}} \quad (\text{S22})$$

in the degenerate case. Note that the energies on the left hand side in the diagonal of the Hamiltonian are the unperturbed eigenvalues E_{n0} while the energies on the right hand side are Kane's modified energies $\tilde{E}_{n\vec{k}}$. This follows from Eq. (S18). The perturbed wave functions at finite \vec{k} for the degenerate case are given by, see Kane 1966, Eq. (4),

$$u_{n\vec{k}} = \sum_i c_{in} u_{i0}, \quad (\text{S23})$$

expressed in terms of the unperturbed wave functions at $\vec{k}=0$.

For specificity and for comparison with the literature, we rewrite equations (S21-S22) by explicitly adding the free-electron term: For a **non-degenerate** band with energy E_{n0} , the correction to the energy up to second order in \vec{k} (and ignoring the linear terms) is, compare Yu & Cardona (2.38),

$$E_{n\vec{k}} = E_{n0} + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0^2} \sum_{i \neq n} \frac{|\langle n0 | \vec{k} \cdot \vec{p} | i0 \rangle|^2}{E_{n0} - E_{i0}}. \quad (\text{S24})$$

For **degenerate** bands, following the eigenvalue problem posed in Eq. (S12), the energies for finite \vec{k} are found by diagonalizing the matrix, see Kane (1966), Eq. (8),

$$\sum_i \left[\left(E_{n0} + \frac{\hbar^2 k^2}{2m_0} \right) \delta_{ni} + \frac{\hbar}{m_0} \langle n0 | \vec{k} \cdot \vec{p} | i0 \rangle \right] c_{ni} = E_{n\vec{k}} c_{nn}. \quad (\text{S25})$$

The eigenvalues of this matrix are the band energies $E_{n\vec{k}}$ near $\vec{k}=0$ from $\vec{k} \cdot \vec{p}$ -theory, including the free-electron energy.

Kane's substitution (S17) only changes the energies, but it does not affect the wave functions. Therefore, Eqs. (S20) and (S23) are valid either way.

It is common to include 3, 4, or 7 bands in the $\vec{k} \cdot \vec{p}$ matrix. A 3×3 -matrix is needed to describe the top valence bands of Si and Ge (see, for example, Kane 1956). With a 4×4 -matrix, one can also include the the lowest conduction band (see Kane 1957). A 7×7 -matrix includes the interaction between the p -bonding valence bands and the p -antibonding conduction bands. This interaction is needed to describe the warping of the valence bands (Dresselhaus, Kip, Kittel). Adding the s -bonding valence band as well completes an 8×8 -matrix. Including higher-energy d -orbitals in the conduction band further increases the size of the matrix (Cardona & Pollak 1966). The rank of this matrix is doubled if the spin degeneracy is included explicitly. For example, a commonly used 14×14 -matrix includes the p -type valence band and the s -type and p -type conduction bands with spin.

S4. APPLICATION TO THE NON-DEGENERATE CONDUCTION BAND IN INSB

We can apply Eq. (S24) to the lowest conduction band of InSb at the Γ -point. The (unperturbed) energy of this band at $\vec{k} = 0$ is called E_0 . The strongest contribution to Eq. (S24) comes from the top valence bands, because they have the smallest energy denominator. The top valence band is threefold degenerate with basis functions X , Y , and Z , see Yu & Cardona (2.50). The only non-vanishing matrix elements in Eq. (S24) are of the form, see Yu & Cardona (2.42),

$$\langle s | p_x | X \rangle = -iP. \quad (\text{S26})$$

The mixed momentum matrix elements $\langle s | p_x | Y \rangle$ etc vanish. Without loss of generality, we can assume that the \vec{k} -vector points along the z -direction. For other directions, one can rotate the coordinate system using Eqs. (6-9) in Kane 1957. Ignoring the contributions from all other "remote" bands, the conduction band energy near the minimum (for small \vec{k}) is found to be

$$\begin{aligned} E_{\text{CB}}(\vec{k}) &= E_0 + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2 k^2}{m_0^2} \frac{P^2}{E_0} = \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left(1 + \frac{2}{m_0} \frac{P^2}{E_0} \right), \end{aligned} \quad (\text{S27})$$

if we neglect the spin-orbit splitting Δ_0 . By selecting the \vec{k} -vector along one of the axes, the sum in Eq. (S24) only has a single term.

If we wish to include the spin-orbit splitting Δ_0 of the valence band maximum, then we need to be careful and use the six basis functions for the valence band given by Eqs. (2.48) and (2.49) in Yu & Cardona instead of the X , Y , Z basis set. See also Eq. (32) in Kane (1966). Still selecting \vec{k} along the z -axis, the sum now has two terms, one for the $j=\frac{3}{2}$ state and one for the $j=\frac{1}{2}$ state. Spin doubles the number of states, but matrix elements between different spin states vanish. The result of Eq. (S24) is

$$E_{\text{CB}}(\vec{k}) = E_0 + \frac{\hbar^2 k^2}{2m_0} \left[1 + \frac{2P^2}{3m_0} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right) \right], \quad (\text{S28})$$

compare Eq. (12) in Kane (1957) or Eq. (5) in Cardona (1963) or Eq. (11) in Emminger (2022). The factor $\frac{1}{3}$ in this equation is derived from the normalization of the basis functions.

We introduce the energy

$$E_P = \frac{2P^2}{m_0}, \quad (\text{S29})$$

which equals about 20 to 25 eV for many common semiconductors. The conduction band energy then becomes

$$E_{\text{CB}}(\vec{k}) = E_0 + \frac{\hbar^2 k^2}{2m_0} \left(1 + \frac{2}{3} \frac{E_P}{E_0} + \frac{1}{3} \frac{E_P}{E_0 + \Delta_0} \right), \quad (\text{S30})$$

We recognize the term in parentheses as the inverse effective mass of the lowest conduction band

$$\frac{1}{m_e^*} = 1 + \frac{2}{3} \frac{E_P}{E_0} + \frac{1}{3} \frac{E_P}{E_0 + \Delta_0} = 1 + \frac{E_P}{3} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right), \quad (\text{S31})$$

compare Emminger (2022), Eq. (11). We stress that this result was obtained using second-order perturbation theory for non-degenerate energy states, keeping only terms up to second order in \vec{k} . The conduction band (S30) turns out to be parabolic, because remote bands and higher order terms in \vec{k} were omitted. Using the values $E_0=0.237$ eV, $\Delta_0=0.81$ eV, and $E_P=23.1$ eV at low temperature

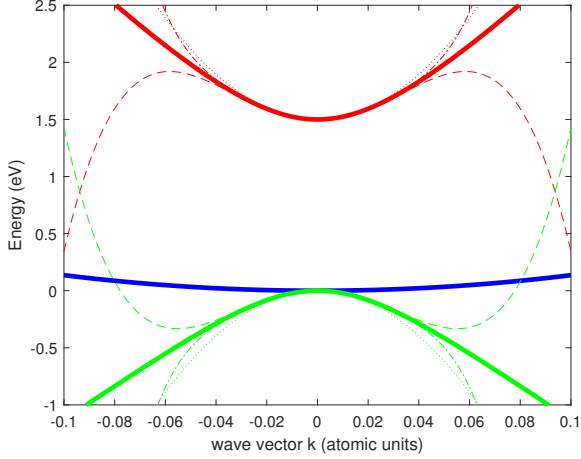


FIG. S1. Band structure of GaAs near the Brillouin zone center calculated from a 4×4 $\vec{k} \cdot \vec{p}$ -model within the approximation for small spin-orbit splitting. The electron, light hole, and heavy hole bands are shown in red, green, and blue, respectively. Thick lines show the dispersion calculated from Eqs. (S35) and (S36) and dotted lines the parabolic approximation. Dashed and dot-dashed lines include the corrections proportional to k^4 and k^6 given by Eq. (S45). The wave vector is shown in atomic units (inverse Bohr radii).

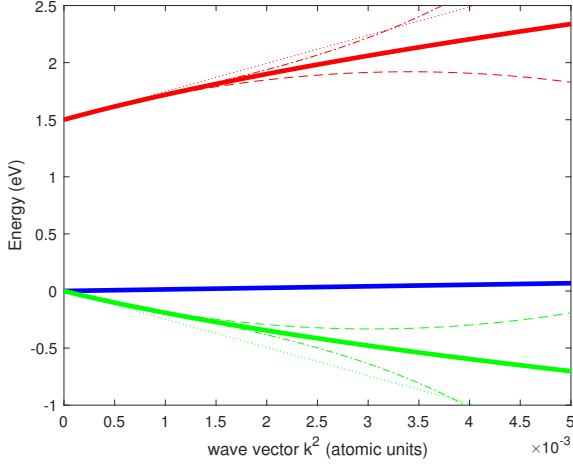


FIG. S2. As Fig. S1, but as a function of the square of the wave vector in atomic units, to better show the deviations from linearity.

(Lawaetz 1971), we find $m_e=0.0136$ for InSb, which compares well to the experimental value of 0.0137. See also Table 2.22 in Yu & Cardona and Sec. IIF in the main manuscript.

S5. BAND ENERGIES FROM 4×4 $\vec{k} \cdot \vec{p}$ -MODEL WITHOUT SPIN-ORBIT COUPLING

To understand the origin of the nonparabolicity of the valence and conduction bands, we start our analysis by evaluating the matrix given in Eq. (S22) with a 4×4 $\vec{k} \cdot \vec{p}$ -model, which includes the three degenerate top valence bands and the lowest conduction band. We first ignore spin-orbit splitting by setting $\Delta_0=0$. (This will be changed later for InSb, where the spin-orbit splitting is large.) We use the basis set $|S\rangle$, $|X\rangle$, $|Y\rangle$, and $|Z\rangle$. The only non-vanishing momentum matrix elements are those given by Eq. (S26). The energy reference level $E_v=0$ is taken to be at the top of the valence band. Without loss of generality, we select the \vec{k} -vector along the x -direction. For other directions, one can rotate the coordinate system using Eqs. (6-9) in Kane 1957.

The $\vec{k} \cdot \vec{p}$ -matrix for the modified energies (S17,S22) is

$$H_0 + \tilde{H}_{\vec{k}} = \begin{pmatrix} E_0 & -\frac{\hbar k}{m_0} iP & 0 & 0 \\ \frac{\hbar k}{m_0} iP & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (\text{S32})$$

where the top row and the left column represent the $|S\rangle$ -state and the bottom row and right column the $|Z\rangle$ -state. The tilde over $H_{\vec{k}}$ indicates that the free-electron kinetic energy has been subtracted as indicated in Eq. (S17). The matrix (S32) is the same as Eq. (4) in Kane (1957), if we set $\Delta_0=0$ and consider that Kane used $|iS\rangle$ as one of the basis functions, which makes the off-diagonal elements real. Compare also Eq. (20) in Kane (1966), which includes many higher-order matrix elements that we have set to zero. But note that Kane (1966) does not introduce the modified energies (S17) with a tilde and instead adds the k^2 term explicitly to the diagonal elements.

This matrix has two eigenvalues $\tilde{E}_{1,2}=0$, compare Eq. (9) in Kane (1957). The other two eigenvalues can be found by solving the characteristic equation

$$\begin{aligned} (E_0 - \tilde{E})(0 - \tilde{E}) - \frac{\hbar^2 k^2 P^2}{m_0^2} &= \\ = \tilde{E}(\tilde{E} - E_0) - \frac{\hbar^2 k^2}{2m_0} E_P &= 0, \end{aligned} \quad (\text{S33})$$

which is equivalent to Kane's (1957) Eq. (10) and Kane's (1966) Eq. (38). The remaining two solutions for the modified energies are

$$\begin{aligned} \tilde{E}_{3,4} &= \frac{E_0}{2} \left[1 \pm \sqrt{1 + \left(\frac{2\hbar k P}{m_0 E_0} \right)^2} \right] \\ &= \frac{E_0}{2} \left(1 \pm \sqrt{1 + 4 \frac{\hbar^2 k^2}{2m_0} \frac{E_P}{E_0^2}} \right), \end{aligned} \quad (\text{S34})$$

which are also given by Kane (1966) in Eq. (43). By adding the k^2 -term back in, see Eq. (S17), we obtain

the four solutions for the energies in the conduction and valence band:

$$E_{1,2} = \frac{\hbar^2 k^2}{2m_0} \quad (\text{S35})$$

$$\begin{aligned} E_{3,4} &= \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left[1 \pm \sqrt{1 + \left(\frac{2\hbar k P}{m_0 E_0} \right)^2} \right] = \\ &= \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left(1 \pm \sqrt{1 + 4 \frac{\hbar^2 k^2 E_P}{2m_0 E_0^2}} \right). \end{aligned} \quad (\text{S36})$$

Equation (S36) can also be applied if the spin-orbit splitting is much smaller than the band gap, see for example Ehrenreich, Phys. Rev. **120**, 1951 (1960), Eq. (1), or Conwell and Vassell, Phys. Rev. **166**, 797 (1968). Eq. (S36) can therefore be used for phosphides, the light arsenides AlAs and GaAs, and with limited accuracy perhaps even for Ge.

We can expand the square root for small values of k , which leads to the parabolic approximation

$$\begin{aligned} E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2 k^2 P^2}{m_0^2 E_0} = \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left(\frac{E_P}{E_0} + 1 \right), \end{aligned} \quad (\text{S37})$$

$$\begin{aligned} E_4 &= \frac{\hbar^2 k^2}{2m_0} - \frac{\hbar^2 k^2 P^2}{m_0^2 E_0} = \\ &= \frac{\hbar^2 k^2}{2m_0} \left(1 - \frac{E_P}{E_0} \right) = -\frac{\hbar^2 k^2}{2m_0} \left(\frac{E_P}{E_0} - 1 \right). \end{aligned} \quad (\text{S38})$$

We need to discuss these results: There are two degenerate bands $E_{1,2}$ with a positive mass of m_0 , which curve upward. Compare Kane 1966, Sec. IV.3.b(2). These bands correspond to the heavy and split-off hole bands. The experimentally observed downward curvature of these bands is therefore caused by the interaction with remote bands at higher energies, which we have ignored in our 4×4 -model, especially the matrix element Q , which describes the interaction between the p -bonding valence band and the p -antibonding conduction band through the DKK ABC parameters, see Emminger (2022), Eqs. (12-20). The dispersion of the band E_3 is the electron in the conduction band and the result is identical to Eq. (S30) for $\Delta_0=0$. Finally, E_4 is the light hole band. It has a negative curvature, since the band gap E_0 is smaller than the momentum matrix element energy E_P . To stress the negative curvature, we have pulled a negative sign in front of the parentheses. This leads to a positive light hole mass.

The effective electron mass from Eq. (S37) is

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{E_0} = \frac{E_P + E_0}{E_0} \approx \frac{E_P}{E_0} \quad (\text{S39})$$

or

$$m_e^* = \frac{E_0}{E_P + E_0} \approx \frac{E_0}{E_P}. \quad (\text{S40})$$

This is the same as Eq. (S31) for $\Delta_0=0$. Similarly, the light hole mass obtained from Eq. (S38) is given by

$$\frac{1}{m_{lh}^*} = \frac{E_P}{E_0} - 1 = \frac{E_P - E_0}{E_0} \quad (\text{S41})$$

or

$$m_{lh}^* = \frac{E_0}{E_P - E_0} \approx \frac{E_0}{E_P} \quad (\text{S42})$$

Within this 4×4 $\vec{k} \cdot \vec{p}$ -model without spin-orbit splitting, the electron and light hole masses (S40) and (S42) are nearly identical.

Using the light hole reduced mass

$$\frac{1}{\mu_{lh}} = \frac{1}{m_{lh}^*} + \frac{1}{m_e^*} = \frac{E_P + E_0}{E_0} + \frac{E_P - E_0}{E_0} = \frac{2E_P}{E_0} \quad (\text{S43})$$

we can write the light hole and electron energies in more compact form

$$E_{3,4} = \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0} \frac{2}{\mu_{lh} E_0}} \right). \quad (\text{S44})$$

It is interesting that the expressions (S44) and (S69) are the same in the small and large spin-orbit splitting approximations, as we shall see later.

The nonparabolicity of the electron and light hole bands is contained in the square-root term in Eq. (S36), if we include higher-order terms in k^2 . By retaining the next-order terms, we obtain

$$\begin{aligned} E_3 &\approx E_0 + \frac{\hbar^2 k^2}{2m_0} \left(1 + \frac{E_P}{E_0} - \frac{\hbar^2 k^2 E_P^2}{2m_0 E_0^3} + 2 \frac{\hbar^4 k^4 E_P^3}{4m_0^2 E_0^5} \right), \\ E_4 &\approx \frac{\hbar^2 k^2}{2m_0} \left(1 - \frac{E_P}{E_0} + \frac{\hbar^2 k^2 E_P^2}{2m_0 E_0^3} - 2 \frac{\hbar^4 k^4 E_P^3}{4m_0^2 E_0^5} \right). \end{aligned} \quad (\text{S45})$$

Since E_P is nearly the same for many common semiconductors, we recognize that nonparabolicity corrections are larger for semiconductors with smaller band gaps.

We can also start the Taylor expansion of the square root with Eq. (S44):

$$\begin{aligned} \tilde{E}_{3,4} &= \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0} \frac{2}{\mu_{lh} E_0}} \right) = \\ &= \frac{E_0}{2} \left[1 \pm 1 \pm \frac{\hbar^2 k^2}{2m_0} \frac{1}{\mu_{lh} E_0} \mp \frac{1}{8} \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{4}{\mu_{lh}^2 E_0^2} \pm \right. \\ &\quad \left. \pm \frac{1}{16} \left(\frac{\hbar^2 k^2}{2m_0} \right)^3 \frac{8}{\mu_{lh}^3 E_0^3} \right]. \end{aligned} \quad (\text{S46})$$

The light hole energy is the expression with the minus sign

$$\begin{aligned} E_4 &= -\frac{\hbar^2 k^2}{2m_0} \left[\left(\frac{1}{2\mu_{lh}} - 1 \right) - \frac{\hbar^2 k^2}{2m_0} \frac{1}{4\mu_{lh}^2 E_0} + \right. \\ &\quad \left. + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{1}{4\mu_{lh}^3 E_0^2} \right] \end{aligned} \quad (\text{S47})$$

We note that the parabolic factor

$$\frac{1}{2\mu_{lh}} - 1 = \frac{E_P}{E_0} - 1 = \frac{E_P - E_0}{E_0} = \frac{1}{m_{lh}^*} \quad (\text{S48})$$

is the inverse light hole mass in the small spin-orbit splitting approximation. The next two higher-order terms in Eq. (S47) agree with those given by Eq. (S45), but written in terms of the reduced light hole mass. This makes the light hole energy

$$E_4 = -\frac{\hbar^2 k^2}{2m_0 m_{lh}^*} \left[1 - \frac{\hbar^2 k^2}{2m_0} \frac{m_{lh}^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{m_{lh}^*}{4\mu_{lh}^3 E_0^2} \right]. \quad (\text{S49})$$

Similarly, the conduction band energy is given by

$$\begin{aligned} E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left[\left(\frac{1}{2\mu_{lh}} + 1 \right) - \frac{\hbar^2 k^2}{2m_0} \frac{1}{4\mu_{lh}^2 E_0} + \right. \\ &\quad \left. + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{1}{4\mu_{lh}^3 E_0^2} \right] \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0 m_e^*} \times \\ &\quad \times \left[1 - \frac{\hbar^2 k^2}{2m_0} \frac{m_e^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{m_e^*}{4\mu_{lh}^3 E_0^2} \right]. \quad (\text{S50}) \end{aligned}$$

The expression

$$\frac{1}{2\mu_{lh}} + 1 = \frac{E_P}{E_0} + 1 = \frac{1}{m_e^*} \quad (\text{S51})$$

is the electron mass.

Figures S1 and S2 show the valence and conduction bands of GaAs calculated from a $4 \times 4 \vec{k} \cdot \vec{p}$ -model without spin orbit-splitting as given by Eqs. (S35) and (S36) and in various approximations.

We need to add a word of caution about the convergence of the series given by Eqs. (S45): For this series to converge, the third term in the series should be smaller than the second one, and the fourth term smaller than the third one. For GaAs, $E_P/E_0 \approx 17$. The approximate convergence criterion is therefore

$$\frac{\hbar^2 k^2}{2m_0} < \frac{E_0}{E_P} E_0 = \frac{E_0}{17} = 0.088 \text{ eV} = 0.003 \text{ Ha}, \quad (\text{S52})$$

where $1 \text{ Ha} = 2 \text{ Ry} = 27.2 \text{ eV}$ is the atomic unit of energy. This implies

$$k^2 < 0.003 \quad \text{or} \quad k < 0.06 \quad (\text{S53})$$

in atomic units, where k is measured in inverse Bohr radii. Inspection of Figs. S1 and S2 indeed shows that the approximations (S45) fail for $k > 0.06$ or $k^2 > 0.004$. For larger wave vectors, the parabolic approximations (S37) and (S38) are actually closer to the exact 4×4 solution (S36) than the nonparabolic corrections (S45). In energy units, the nonparabolic corrections (S45) are only beneficial up to about 0.4 eV above the band gap for GaAs.

S6. BAND ENERGIES FROM $8 \times 8 \vec{k} \cdot \vec{p}$ -MODEL WITH SPIN-ORBIT COUPLING

We now proceed to discuss how a finite spin-orbit splitting Δ_0 affects the band structure in a $4 \times 4 \vec{k} \cdot \vec{p}$ model. Following Kane (1957), we use a basis set $|S\rangle$, $|X \pm iY\rangle/\sqrt{2}$, and $|Z\rangle$ and assume that the \vec{k} vector is along the z -direction. For other directions, one can rotate the coordinate system using Eqs. (6-9) in Kane 1957. The matrix is doubled to include the two spin states. In addition to the $\vec{k} \cdot \vec{p}$ perturbation Hamiltonian $H_{\vec{k}}$, we also include the k -independent spin-orbit interaction as a second perturbation. This spin-orbit Hamiltonian is given in Eq. (1) of Kane (1957) and also by Yu & Cardona Eq. (2.45a)

$$H_{\text{SO}} = \frac{\hbar}{4c^2 m_0^2} \left(\vec{\nabla} V \times \vec{p} \right) \cdot \vec{\sigma}, \quad (\text{S54})$$

where $\vec{\sigma}$ is a vector formed by the three Pauli matrices and c is the speed of light. There is also a k -dependent spin-orbit interaction, which is neglected. Δ_0 is the matrix element of the spin-orbit Hamiltonian (Kane 1957)

$$\Delta_0 = \frac{3\hbar i}{4m_0^2 c^2} \left\langle X \left| \frac{\partial V}{\partial x} p_y - \frac{\partial V}{\partial y} p_x \right| Y \right\rangle. \quad (\text{S55})$$

The resulting $8 \times 8 \vec{k} \cdot \vec{p}$ -matrix consists of two identical 4×4 blocks of the form

$$H_0 + \tilde{H}_{\vec{k}} + H_{\text{SO}} = \begin{pmatrix} \tilde{H}_4 & 0 \\ 0 & \tilde{H}_4 \end{pmatrix}, \quad (\text{S56})$$

where the 4×4 block is given by, see Kane (1957) Eq. (4) or Kane (1966) Eq. (36)

$$\tilde{H}_{\vec{k}} = \begin{pmatrix} E_0 & 0 & -\frac{\hbar k}{m_0} iP & 0 \\ 0 & -\frac{2\Delta_0}{3} & \frac{\sqrt{2}\Delta_0}{3} & 0 \\ \frac{\hbar k}{m_0} iP & \frac{\sqrt{2}\Delta_0}{3} & -\frac{\Delta_0}{3} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (\text{S57})$$

The tilde over $H_{\vec{k}}$ indicates that the free-electron kinetic energy has been subtracted as indicated in Eq. (S17). A different Hamiltonian (with the energy eigenvalues at $k=0$ on the diagonal) is obtained if one chooses the wave functions given by Eqs. (2.48-2.49) by Yu and Cardona.

The matrix (S57) has one obvious eigenvalue $\tilde{E}_1 = 0$ (which we can call the heavy hole) and the other three are determined from the characteristic polynomial of the upper 3×3 block

$$\begin{aligned} &\left(E_0 - \tilde{E} \right) \left(-\frac{2\Delta_0}{3} - \tilde{E} \right) \left(-\frac{\Delta_0}{3} - \tilde{E} \right) - \\ &- \frac{\hbar^2 k^2}{m_0^2} P^2 \left(-\frac{2\Delta_0}{3} - \tilde{E} \right) - \frac{2}{9} \Delta_0^2 \left(E_0 - \tilde{E} \right) = 0. \quad (\text{S58}) \end{aligned}$$

The first and third term in Eq. (S58) can be combined to cancel the quadratic spin-orbit splitting term. By changing the overall sign, the characteristic equation becomes

a cubic equation of the form

$$\tilde{E} \left(\tilde{E} - E_0 \right) \left(\tilde{E} + \Delta_0 \right) - \frac{\hbar^2 k^2 E_P}{2m_0} \left(\tilde{E} + \frac{2\Delta_0}{3} \right) = 0, \quad (\text{S59})$$

which was given by Kane (1957) in Eq. (10) and by Kane (1966) in Eq. (38). See also Bartoli (PRB 27, 2248, 1983) in more common recent notation identical to Eq. (S59). This becomes equal to the earlier characteristic equation (S33) if we set the spin-orbit splitting $\Delta_0=0$. Kane (1966) calls Eq. (S59) the small-gap approximation, because it ignores the interactions with other "remote" bands, for example the p-antibonding upper conduction band, the s-bonding lower valence band, or d-orbitals.

This characteristic equation (S59) is a cubic equation, which can be solved analytically with modern computers. We will show the results below, but first we discuss the approximation that the spin-orbit splitting is much larger than the band gap, which is appropriate for InSb.

S7. APPROXIMATION FOR LARGE SPIN-ORBIT SPLITTING

If the spin-orbit splitting Δ_0 is large, then we can make the approximation $\tilde{E} + \Delta_0 \approx \Delta_0$. The secular equation (S59) then becomes quadratic and reduces to

$$\tilde{E} \left(\tilde{E} - E_0 \right) - \frac{2\hbar^2 k^2}{3} E_P = 0 \quad (\text{S60})$$

because the Δ_0 term can be cancelled. This characteristic equation is the same as in Eq. (S33) except for a factor $\frac{2}{3}$. The conduction and light hole bands are then described by an expression similar to Eq. (S36), except for the additional factor $\frac{2}{3}$, and are given by

$$\begin{aligned} E_{3,4} &= \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{8\hbar^2 k^2 P^2}{3m_0^2 E_0^2}} \right) = \\ &= \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{8\hbar^2 k^2 E_P}{3} \frac{E_0}{2m_0 E_0^2}} \right), \end{aligned} \quad (\text{S61})$$

compare Eq. (13) in Kane (1957) or Eq. (50) in Kane (1966). The heavy hole band in the 8×8 approximation (which ignores interactions with the p-antibonding conduction band) still has the wrong curvature given by Eq. (S35). The split-off hole band E_2 is not included in this approximation.

We can expand the square root for small values of k , which leads to the parabolic approximation

$$\begin{aligned} E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0} + \frac{2\hbar^2 k^2 P^2}{3m_0^2 E_0} = \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left(\frac{2E_P}{3E_0} + 1 \right), \end{aligned} \quad (\text{S62})$$

$$E_4 = \frac{\hbar^2 k^2}{2m_0} - \frac{2\hbar^2 k^2 P^2}{3m_0^2 E_0} = -\frac{\hbar^2 k^2}{2m_0} \left(\frac{2E_P}{3E_0} - 1 \right) \quad (\text{S63})$$

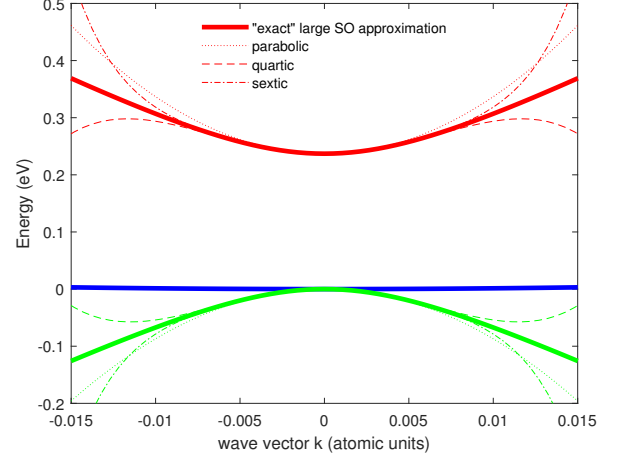


FIG. S3. Band structure of InSb near the Brillouin zone center calculated from a $4 \times 4 \vec{k} \cdot \vec{p}$ -model within the approximation for large spin-orbit splitting. The electron, light hole, and heavy hole bands are shown in red, green, and blue, respectively. Thick lines show the dispersion calculated from Eqs. (S61) and dotted lines the parabolic approximation (S62) and (S63). Dashed and dot-dashed lines include the corrections proportional to k^4 and k^6 given by Eq. (S70) and (S71). The wave vector is shown in atomic units (inverse Bohr radii).

The effective electron mass in this approximation for large spin-orbit splitting is

$$\frac{1}{m_e^*} = 1 + \frac{2E_P}{3E_0} = \frac{2E_P + 3E_0}{3E_0} \approx \frac{2E_P}{3E_0} \quad (\text{S64})$$

or

$$m_e^* = \frac{3E_0}{2E_P + 3E_0} \approx \frac{3E_0}{2E_P}. \quad (\text{S65})$$

as expected from Eq. (S31) with a large spin-orbit splitting Δ_0 . Similarly, the light hole mass for large spin-orbit splittings is

$$\frac{1}{m_{lh}^*} = - \left(1 - \frac{2E_P}{3E_0} \right) = \frac{2E_P - 3E_0}{3E_0} \approx \frac{2E_P}{3E_0} \quad (\text{S66})$$

or

$$m_{lh}^* = \frac{3E_0}{2E_P - 3E_0} \approx \frac{3E_0}{2E_P}, \quad (\text{S67})$$

if we keep in mind that the overall minus sign is needed to ensure a negative curvature for a positive hole mass. This also agrees with Eq. (S31) in the limit of large Δ_0 . The electron and light hole masses are again nearly identical for $E_0 \ll E_P$.

Using the light hole reduced mass

$$\frac{1}{\mu_{lh}} = \frac{1}{m_{lh}} + \frac{1}{m_e} = \frac{2E_P - 3E_0}{3E_0} + \frac{2E_P + 3E_0}{3E_0} = \frac{4E_P}{3E_0}, \quad (\text{S68})$$

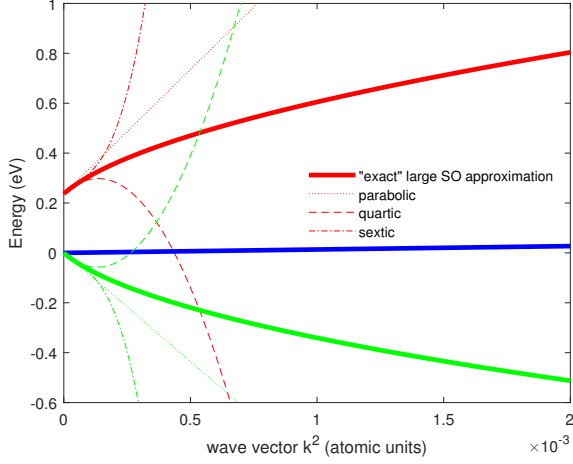


FIG. S4. As Fig. S3, but as a function of the square of the wave vector in atomic units. For comparison, with Fig. 1 in Kane (1957), the range of wave vectors was expanded.

we can write the light hole and electron energies in more compact form

$$E_{3,4} = \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0} \frac{2}{\mu_{lh} E_0}} \right). \quad (\text{S69})$$

It is interesting that the expressions (S44) and (S69) are the same in the small and large spin-orbit splitting approximations.

The nonparabolicity of the electron and light hole bands is contained in the square-root term in Eq. (S61), if we include higher-order terms in k^2 . By retaining the next-order terms, we obtain

$$E_3 = E_0 + \frac{\hbar^2 k^2}{2m_0} \times \left(1 + \frac{2 E_P}{3 E_0} - \frac{4 \hbar^2 k^2 E_P^2}{9 2m_0 E_0^3} + 2 \frac{8 \hbar^4 k^4 E_P^3}{27 4m_0^2 E_0^5} \right), \quad (\text{S70})$$

$$E_4 = \frac{\hbar^2 k^2}{2m_0} \times \left(1 - \frac{2 E_P}{3 E_0} + \frac{4 \hbar^2 k^2 E_P^2}{9 2m_0 E_0^3} - 2 \frac{8 \hbar^4 k^4 E_P^3}{27 4m_0^2 E_0^5} \right). \quad (\text{S71})$$

These equations are similar to those for small spin-orbit splittings given by Eqs. (S45) except for the powers of $\frac{2}{3}$ that appear as factors.

We can also start the Taylor expansion of the square root with Eq. (S69):

$$\begin{aligned} \tilde{E}_{3,4} &= \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0} \frac{2}{\mu_{lh} E_0}} \right) = \\ &= \frac{E_0}{2} \left[1 \pm 1 \pm \frac{\hbar^2 k^2}{2m_0} \frac{1}{\mu_{lh} E_0} \mp \frac{1}{8} \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{4}{\mu_{lh}^2 E_0^2} \pm \right. \end{aligned}$$

$$\left. \pm \frac{1}{16} \left(\frac{\hbar^2 k^2}{2m_0} \right)^3 \frac{8}{\mu_{lh}^3 E_0^3} \right]. \quad (\text{S72})$$

The light hole energy is the expression with the minus sign

$$E_4 = -\frac{\hbar^2 k^2}{2m_0} \left[\left(\frac{1}{2\mu_{lh}} - 1 \right) - \frac{\hbar^2 k^2}{2m_0} \frac{1}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{1}{4\mu_{lh}^3 E_0^2} \right] \quad (\text{S73})$$

We note that the parabolic factor

$$\frac{1}{2\mu_{lh}} - 1 = \frac{2E_P}{3E_0} - 1 = \frac{2E_P - 3E_0}{3E_0} = \frac{1}{m_{lh}^*} \quad (\text{S74})$$

is the inverse light hole mass in the large spin-orbit splitting approximation. The next two higher-order terms in Eq. (S73) agree with those given by Eq. (S71), but written in terms of the reduced light hole mass. This makes the light hole energy

$$E_4 = -\frac{\hbar^2 k^2}{2m_0 m_{lh}^*} \left[1 - \frac{\hbar^2 k^2}{2m_0} \frac{m_{lh}^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{m_{lh}^*}{4\mu_{lh}^3 E_0^2} \right], \quad (\text{S75})$$

just like Eq. (S49) in the case of small spin-orbit splittings.

Similarly, the conduction band energy is given by

$$\begin{aligned} E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left[\left(\frac{1}{2\mu_{lh}} + 1 \right) - \frac{\hbar^2 k^2}{2m_0} \frac{1}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{1}{4\mu_{lh}^3 E_0^2} \right] \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0 m_e^*} \times \\ &\times \left[1 - \frac{\hbar^2 k^2}{2m_0} \frac{m_e^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{m_e^*}{4\mu_{lh}^3 E_0^2} \right], \quad (\text{S76}) \end{aligned}$$

just like Eq. (S47) in the case of small spin-orbit splittings. The expression

$$\frac{1}{2\mu_{lh}} + 1 = \frac{2E_P}{3E_0} + 1 = \frac{1}{m_e^*} \quad (\text{S77})$$

is the electron mass, just like Eq. (S51) in the small spin-orbit splitting approximation.

Kane (1966) Eq. (50) and (1957) Eq. (13) also lists the dispersion of the split-off hole

$$\begin{aligned} E_2 &= -\Delta_0 + \frac{\hbar^2 k^2}{2m_0} - \frac{\hbar^2 k^2 P^2}{3m_0^2 (E_0 + \Delta_0)} = \\ &= -\Delta_0 + \frac{\hbar^2 k^2}{2m_0} - \frac{\hbar^2 k^2 E_P}{2m_0 3(E_0 + \Delta_0)} \\ &= -\Delta_0 - \frac{\hbar^2 k^2}{2m_0} \left[\frac{E_P}{3(E_0 + \Delta_0)} - 1 \right] \quad (\text{S78}) \end{aligned}$$

See also Yu & Cardona Eq. (2.59). We will derive this expression later in the context of the perturbative solution to the cubic characteristic equation in Sec. S9. The inverse split-off mass is therefore

$$\frac{1}{m_{\text{so}}^*} = \frac{E_P}{3(E_0 + \Delta_0)} - 1. \quad (\text{S79})$$

In principle, these expressions are expected to be quite accurate for InSb, where the condition $\Delta_0 \gg E_0$ is satisfied, except for the contributions of "remote" bands. (For α -Sn, on the other hand, \bar{E}_0 is about half of Δ_0 and therefore the approximation for large spin-orbit splitting will be less accurate.) However, since $E_P/E_0 \approx 100$ for InSb, the range of applicability of Eqs. (S70) and (S71) is even more limited than for GaAs. Using the same convergence criterion (S52) as for GaAs, we expect the nonparabolicity corrections to converge only if

$$k^2 < 0.1 \times 10^{-3} \quad \text{or} \quad k < 0.01 \quad (\text{S80})$$

for InSb (in atomic units for k). In energy units, the nonparabolic corrections (S70) and (S71) are not useful more than 70 meV from the band edge.

Figure S3 shows the top valence bands and the lowest conduction band of InSb in the center of the Brillouin zone. The dotted lines show the parabolic electron and light hole bands. The electron and light hole masses were calculated with Eq. (S31) using the parameters $E_0=0.237$ eV, $\Delta_0=0.81$ eV, and $E_P=23.1$ eV given by Lawaetz 1971. The thick solid lines show the nonparabolic bands for the electron, heavy hole and light hole bands calculated using Eqs. (S61) within the large spin-orbit splitting approximation. The nonparabolic corrections given by Eqs. (S70) and (S71) are only applicable for very small wave vectors, as shown by the dashed and dot-dashed lines.

For comparison with Fig. 1 in Kane (1957), we also show the energies near $k=0$ on an expanded wave vector scale as a function of the square of the wave vector in atomic units, see Fig. S4. Despite neglecting the influence of "remote" bands, the agreement with Kane (1957) is quite good.

Additional improvements to the electron and hole dispersions for large spin-orbit splittings were given by Kane (1966) in Eqs. (52-56). Also, corrections to the electron and hole dispersions were given by Earnest Johnson (Semiconductors and Semimetals, Vol. 3, 1967), Eqs. (41-44) without an explanation.

S8. VIETA'S SOLUTION TO THE CUBIC CHARACTERISTIC EQUATION

The eigenvalues of the 4×4 Hamiltonian (S57) are given by the characteristic equation (S59)

$$\tilde{E} \left(\tilde{E} - E_0 \right) \left(\tilde{E} + \Delta_0 \right) - \frac{\hbar^2 k^2}{2m_0} E_P \left(\tilde{E} + \frac{2\Delta_0}{3} \right) = 0. \quad (\text{S81})$$

With the substitution

$$A^2 = \frac{1}{3} \frac{\hbar^2 k^2}{2m_0} E_P, \quad (\text{S82})$$

this can also be written as

$$\tilde{E} \left(\tilde{E} - E_0 \right) \left(\tilde{E} + \Delta_0 \right) - 3A^2 \left(\tilde{E} + \frac{2\Delta_0}{3} \right) = 0. \quad (\text{S83})$$

We collect the powers of \tilde{E} to find the cubic equation

$$\tilde{E}^3 + \tilde{E}^2 (\Delta_0 - E_0) - \tilde{E} (E_0 \Delta_0 + 3A^2) - 2A^2 \Delta_0 = 0. \quad (\text{S84})$$

This equation is of the form

$$ax^3 + bx^2 + cx + d = 0 \quad (\text{S85})$$

with the coefficients

$$x = \tilde{E} = E - \frac{\hbar^2 k^2}{2m_0}, \quad (\text{S86})$$

$$a = 1, \quad (\text{S87})$$

$$b = \Delta_0 - E_0, \quad (\text{S88})$$

$$c = -E_0 \Delta_0 - 3A^2, \quad (\text{S89})$$

$$d = -2A^2 \Delta_0. \quad (\text{S90})$$

This general cubic equation can be reduced to a depressed cubic equation (see Wikipedia)

$$t^3 + pt + q = 0 \quad (\text{S91})$$

with the substitution

$$t = x + \frac{b}{3a} = x + \frac{1}{3} (\Delta_0 - E_0), \quad (\text{S92})$$

$$p = \frac{3ac - b^2}{3a^2} = -E_0 \Delta_0 - 3A^2 - \frac{1}{3} (\Delta_0 - E_0)^2, \quad (\text{S93})$$

$$q = \frac{2b^3 - 9abc + 27a^2d}{27a^3} = \frac{2}{27} (\Delta_0 - E_0)^3 + \frac{1}{3} (\Delta_0 - E_0) (E_0 \Delta_0 + 3A^2) - 2A^2 \Delta_0. \quad (\text{S94})$$

The three solutions to the depressed cubic equation (due to Vieta, see Wikipedia) are given by

$$t_n = 2\sqrt{\frac{-p}{3}} \cos \left[\frac{1}{3} \arccos \left(\frac{3q}{2p} \sqrt{\frac{-3}{p}} \right) - \frac{2\pi n}{3} \right] \quad (\text{S95})$$

for $n=0, 1, 2$. We note that p is negative. Therefore, the arguments under the square roots in Eq. (S95) are real.

Using these expressions, the exact solution to the cubic characteristic equation for the 4×4 $\vec{k} \cdot \vec{p}$ Hamiltonian can be programmed in MATLAB. Results are shown in Figs. S5 and S6 for GaAs and InSb together with the appropriate small or large spin-orbit approximation. The corrections to the conduction band are moderate, but large for the light hole band in GaAs, which strongly

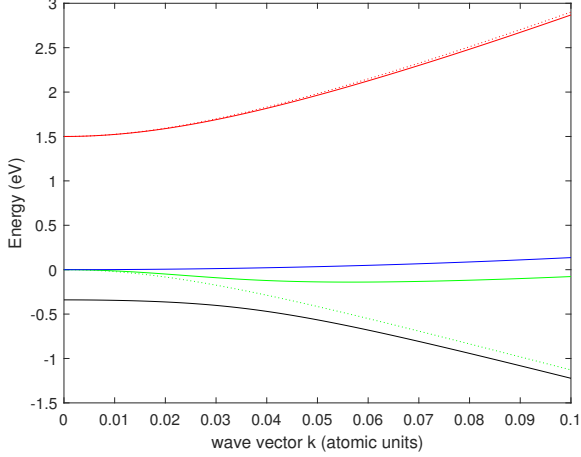


FIG. S5. Heavy (blue), light (green), and split-off hole (black) and electron bands (red) of GaAs near $k=0$ calculated from the cubic characteristic equation (S59) of the $4 \times 4 \vec{k} \cdot \vec{p}$ Kane Hamiltonian (S57) for finite spin-orbit splitting (solid) and from the small spin-orbit approximation (S36) (dotted).

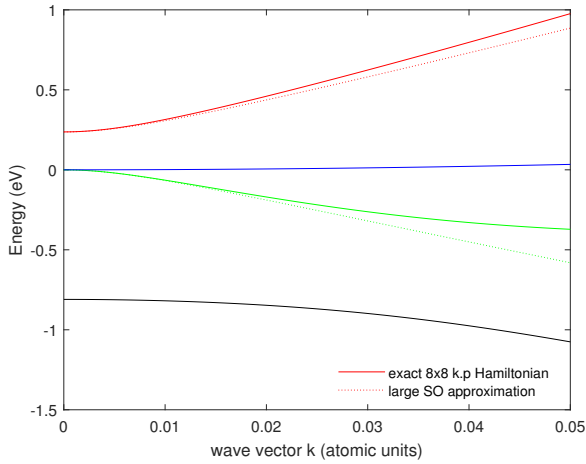


FIG. S6. Heavy (blue), light (green), and split-off hole (black) and electron bands (red) of InSb near $k=0$ calculated from the cubic characteristic equation (S59) of the $4 \times 4 \vec{k} \cdot \vec{p}$ Kane Hamiltonian (S57) for finite spin-orbit splitting (solid) and from the large spin-orbit approximation (S61) (dotted).

couples to the split-off hole band, as mentioned by Bartoli (1983).

Figure S7 shows the inverted band structure of α -tin with the Γ_7^- s-antibonding valence band in green and the p-bonding Γ_8^+ conduction band in red, calculated from the $4 \times 4 \vec{k} \cdot \vec{p}$ Kane Hamiltonian (S57) (solid). The deviations from the parabolic bands (dotted) are even stronger for α -tin than for GaAs and InSb, presumably because of the strong interaction of these bands with the Γ_7^+ split-off hole band. The Γ_7^- and Γ_8^+ valence bands are nearly parallel across a large volume of the Brillouin zone, giv-

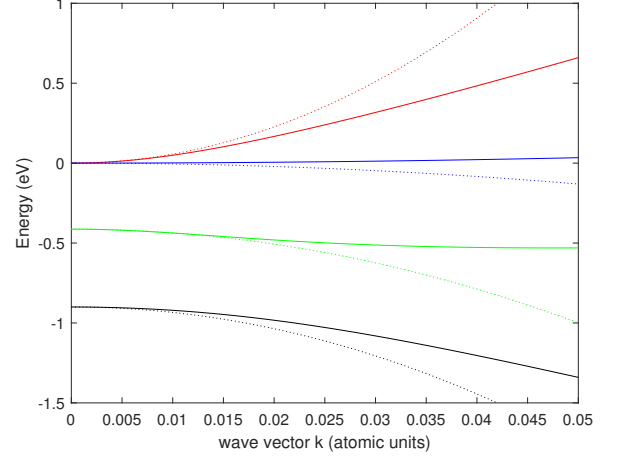


FIG. S7. Γ_8^+ conduction (red) and valence (blue) bands, Γ_7^- valence band (green), and Γ_7^+ split-off band for α -Sn calculated from the $4 \times 4 \vec{k} \cdot \vec{p}$ Kane Hamiltonian (S57) (solid) and the parabolic bands (dotted) with effective masses given by Carrasco (2018).

ing rise to strong allowed intravalence band transitions at the inverted \bar{E}_0 band gap in p-type α -tin or at room temperature.

S9. PERTURBATIVE SOLUTION OF THE 4×4 CHARACTERISTIC EQUATION

The approximations for small and large spin-orbit splittings derived in Secs. S5 and S6 resulted in expressions (S40) and (S65) for the effective electron mass. Neither of these are satisfactory, because they do not yield the electron mass (S31) calculated using non-degenerate perturbation theory. Instead, we pursue a perturbative approach to solve the cubic characteristic equation

$$\tilde{E}^3 + \tilde{E}^2 (\Delta_0 - E_0) - \tilde{E} (E_0 \Delta_0 + 3A^2) - 2A^2 \Delta_0 = 0 \quad (\text{S96})$$

for small wave vectors (but arbitrary values of E_0 and Δ_0), where A was defined in Eq. (S82). With the substitution

$$\epsilon = \frac{3A^2}{E_0 \Delta_0} = \frac{\hbar^2 k^2}{2m_0} \frac{E_P}{\Delta_0 E_0}, \quad (\text{S97})$$

the characteristic equation becomes

$$\tilde{E}^3 + \tilde{E}^2 (\Delta_0 - E_0) - \tilde{E} E_0 \Delta_0 (1 + \epsilon) - \frac{2}{3} \epsilon E_0 \Delta_0^2 = 0, \quad (\text{S98})$$

For small wave vectors (and thus small values of ϵ), we write the modified energy \tilde{E} as a Taylor series in powers of ϵ , up to second order

$$\tilde{E} = \lambda_0 + \lambda_1 \epsilon + \lambda_2 \epsilon^2 \quad (\text{S99})$$

and substitute Eq. (S99) into Eq. (S98). With the intermediate expressions

$$\tilde{E}^2 \approx \lambda_0^2 + 2\lambda_0\lambda_1\epsilon + (\lambda_1^2 + 2\lambda_0\lambda_2)\epsilon^2, \quad (\text{S100})$$

$$\tilde{E}^3 \approx \lambda_0^3 + 3\lambda_0^2\lambda_1\epsilon + 3(\lambda_0^2\lambda_2 + \lambda_0\lambda_1^2)\epsilon^2, \quad (\text{S101})$$

the characteristic equation (S98) becomes

$$\begin{aligned} & \lambda_0^3 + 3\lambda_0^2\lambda_1\epsilon + 3(\lambda_0^2\lambda_2 + \lambda_0\lambda_1^2)\epsilon^2 + \\ & + [\lambda_0^2 + 2\lambda_0\lambda_1\epsilon + (\lambda_1^2 + 2\lambda_0\lambda_2)\epsilon^2](\Delta_0 - E_0) - \\ & - (\lambda_0 + \lambda_1\epsilon + \lambda_2\epsilon^2)E_0\Delta_0(1 + \epsilon) - \frac{2}{3}\epsilon E_0\Delta_0^2 = 0 \end{aligned} \quad (\text{S102})$$

This characteristic equation (S102) must be valid for all orders of ϵ . This yields three equations for the Taylor series coefficients in Eq. (S99).

$$\lambda_0^3 + \lambda_0^2(\Delta_0 - E_0) - \lambda_0 E_0 \Delta_0 = 0, \quad (\text{S103})$$

$$\begin{aligned} & 3\lambda_0^2\lambda_1 + 2\lambda_0\lambda_1(\Delta_0 - E_0) - \\ & - (\lambda_0 + \lambda_1)E_0\Delta_0 - \frac{2}{3}E_0\Delta_0^2 = 0, \end{aligned} \quad (\text{S104})$$

$$\begin{aligned} & 3(\lambda_0^2\lambda_2 + \lambda_0\lambda_1^2) + \\ & + (\lambda_1^2 + 2\lambda_0\lambda_2)(\Delta_0 - E_0) - \\ & - (\lambda_1 + \lambda_2)E_0\Delta_0 = 0. \end{aligned} \quad (\text{S105})$$

If we are interested in the dispersion of the light hole bands, then we know that $\lambda_0=0$. Equation (S104) therefore yields

$$\lambda_1 = -\frac{2}{3}\Delta_0 \quad (\text{S106})$$

and Eq. (S105) yields

$$\lambda_2 = \frac{4}{9}\frac{\Delta_0^2}{E_0} + \frac{2}{9}\Delta_0 = \frac{2}{9}\Delta_0 \left(1 + \frac{2\Delta_0}{E_0}\right). \quad (\text{S107})$$

We insert this into Eq. (S99) and find the modified light hole energy to be

$$\tilde{E}_{\text{lh}} = -\frac{2}{3}\frac{\hbar^2 k^2}{2m_0}\frac{E_P}{E_0} + \frac{2}{9}\left(1 + \frac{2\Delta_0}{E_0}\right)\left(\frac{\hbar^2 k^2}{2m_0}\right)^2 \frac{E_P^2}{\Delta_0 E_0^2}. \quad (\text{S108})$$

The light hole energy E_{lh} is determined from Eq. (S17) and turns out to be

$$E_{\text{lh}} = \frac{\hbar^2 k^2}{2m_0} \left[1 - \frac{2}{3}\frac{E_P}{E_0} + \frac{2}{9}\frac{E_P^2}{\Delta_0 E_0^2} \left(1 + \frac{2\Delta_0}{E_0}\right) \frac{\hbar^2 k^2}{2m_0} \right]. \quad (\text{S109})$$

The first (quadratic) term in this expression yields the light hole mass

$$\frac{1}{m_{\text{lh}}^*} = \frac{2}{3}\frac{E_P}{E_0} - 1 = \frac{2E_P - 3E_0}{3E_0}. \quad (\text{S110})$$

This is the same expression as in Eq. (S66) for the large spin-orbit approximation. For the small spin-orbit splitting approximation, Eq. (S42) yields a different result,

because the strong interaction with the split-off band has been neglected. The second (quartic) term in Eq. (S109) approaches the lowest-order nonparabolic correction in the limit $\Delta_0 \gg E_0$ given by Eq. (S71), but offers a modification for a finite spin-orbit splitting. As shown in Fig. S8, the perturbative solution (S109) to the cubic characteristic equation actually provides a poorer approximation to the full solution in Sec. S8 than the lowest-order nonparabolic correction in Eq. (S71) in the approximation for large spin-orbit splittings.

We find the dispersion of the split-off hole band by setting $\lambda_0 = -\Delta_0$. For this band, Eq. (S104) yields

$$\lambda_1 = -\frac{1}{3}\frac{E_0\Delta_0}{E_0 + \Delta_0} \quad (\text{S111})$$

and Eq. (S105) yields

$$\lambda_2 = -\frac{E_0^2\Delta_0(\Delta_0 + 2E_0)}{9(E_0 + \Delta_0)^3}. \quad (\text{S112})$$

We insert these expressions into Eq. (S99) and find the split-off hole energy

$$\begin{aligned} E_{\text{so}} = & -\Delta_0 - \frac{\hbar^2 k^2}{2m_0} \left[\frac{E_P}{3(E_0 + \Delta_0)} - 1 \right] - \\ & - \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{E_P^2}{\Delta_0} \frac{\Delta_0 + 2E_0}{9(E_0 + \Delta_0)^3}. \end{aligned} \quad (\text{S113})$$

The quadratic term in Eq. (S113) contains the inverse effective mass of the split-off hole (in square brackets), which we have already mentioned in Eq. (S79).

Finally, we find the dispersion of the conduction band by setting $\lambda_0 = E_0$. For this band, Eq. (S104) yields

$$\lambda_1 = \frac{\Delta_0(E_0 + \frac{2}{3}\Delta_0)}{E_0 + \Delta_0} = \frac{E_0\Delta_0}{3} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right) \quad (\text{S114})$$

and the parabolic expression for the conduction band is therefore

$$E_e = E_0 + \frac{\hbar^2 k^2}{2m_0} \left[1 + \frac{E_P}{E_0} \frac{E_0 + \frac{2}{3}\Delta_0}{E_0 + \Delta_0} \right] \quad (\text{S115})$$

$$= E_0 + \frac{\hbar^2 k^2}{2m_0} \left[1 + \frac{E_P}{3} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right) \right] \quad (\text{S116})$$

$$= E_0 + \frac{\hbar^2 k^2}{2m_0 m_e}. \quad (\text{S117})$$

The expression in square brackets is the inverse effective electron mass

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{3} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right) \quad (\text{S118})$$

given already in (S31) as found from non-degenerate perturbation theory. It also agrees with Eq. (S40) in the limit of $\Delta_0=0$ and with Eq. (S65) in the limit of large

spin-orbit splittings. The inverse reduced light hole mass in this approximation is

$$\frac{1}{\mu_{lh}} = \frac{1}{m_{lh}^*} + \frac{1}{m_e^*} = \frac{E_P}{3} \left(\frac{4}{E_0} + \frac{1}{E_0 + \Delta_0} \right). \quad (\text{S119})$$

Equation (S105) yields

$$\lambda_2 = -\frac{\lambda_1 [\lambda_1 (2E_0 + \Delta_0) - E_0 \Delta_0]}{E_0 (E_0 + \Delta_0)}. \quad (\text{S120})$$

We note that

$$\lambda_1 \epsilon = \left(\frac{1}{m_e} - 1 \right) \frac{\hbar^2 k^2}{2m_0}. \quad (\text{S121})$$

and therefore

$$\begin{aligned} \lambda_2 \epsilon^2 &= -\frac{\lambda_1 \epsilon [\lambda_1 \epsilon (2E_0 + \Delta_0) - E_0 \Delta_0 \epsilon]}{E_0 (E_0 + \Delta_0)} = \\ &= -\left(\frac{1}{m_e} - 1 \right)^2 \frac{2E_0 + \Delta_0 - \frac{E_P}{\frac{1}{m_e} - 1}}{E_0 (E_0 + \Delta_0)} \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \end{aligned}$$

The middle factor can be simplified using Eq. (S31) as follows

$$\begin{aligned} &\frac{2E_0 + \Delta_0 - \frac{E_P}{\frac{1}{m_e} - 1}}{E_0 (E_0 + \Delta_0)} = \\ &= \frac{2E_0 + \Delta_0 - \frac{E_P}{\frac{3E_0 + 2\Delta_0}{3E_0 + 2\Delta_0}}}{E_0 (E_0 + \Delta_0)} = \\ &= \frac{2E_0 + \Delta_0 - 3 \frac{E_0 (E_0 + \Delta_0)}{3E_0 + 2\Delta_0}}{E_0 (E_0 + \Delta_0)} = \\ &= \frac{3E_0^2 + 2E_0 \Delta_0 + 2E_0 \Delta_0 + 2\Delta_0^2}{E_0 (E_0 + \Delta_0) (3E_0 + 2\Delta_0)} = \\ &= \frac{E_0 (3E_0 + 2\Delta_0) + 2\Delta_0 (E_0 + \Delta_0)}{E_0 (E_0 + \Delta_0) (3E_0 + 2\Delta_0)} = \\ &= \frac{1}{E_0 + \Delta_0} + \frac{2\Delta_0}{E_0} \frac{1}{3E_0 + 2\Delta_0}. \quad (\text{S122}) \end{aligned}$$

We insert these expressions into Eq. (S99) and find the conduction band electron energy

$$\begin{aligned} E_e &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left[1 + \frac{E_P}{E_0} \frac{E_0 + \frac{2}{3}\Delta_0}{E_0 + \Delta_0} \right] - \\ &\quad - \left(\frac{1}{m_e} - 1 \right)^2 \left(\frac{1}{E_0 + \Delta_0} + \frac{2\Delta_0}{E_0} \frac{1}{3E_0 + 2\Delta_0} \right) \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \\ &= E_0 + \frac{\hbar^2 k^2}{2m_0} \left[1 + \frac{E_P}{3} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right) \right] - \quad (\text{S123}) \\ &\quad - \left(\frac{1}{m_e} - 1 \right)^2 \left(\frac{1}{E_0 + \Delta_0} + \frac{2\Delta_0}{E_0} \frac{1}{3E_0 + 2\Delta_0} \right) \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \end{aligned}$$

The result of this approximation (S123) is shown in Fig. S8. It provides about the same accuracy as the lowest-order nonparabolic correction for large spin-orbit splittings given by Eq. (S71).

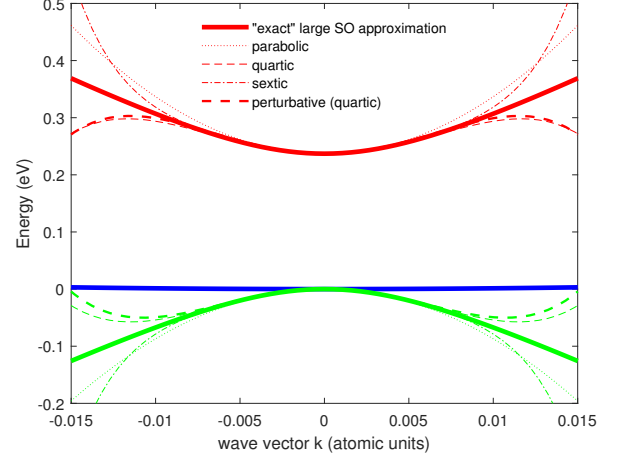


FIG. S8. As Fig. S6 for InSb, but we have added thick dashed lines to indicate the light hole and electron dispersions from the perturbative solutions to the cubic characteristic equation, as given by Eqs. (S109) and (S123).

S10. HEURISTIC INTERPOLATION BETWEEN SMALL AND LARGE SPIN-ORBIT SPLITTINGS

As we have seen previously during the derivation of the light hole and electron energies for small and large spin-orbit splittings, both approximations yield the same expressions given by Eqs. (S44) and (S69)

$$E_{3,4} = \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{\hbar^2 k^2}{2m_0} \frac{2}{\mu_{lh} E_0}} \right), \quad (\text{S124})$$

but the definitions of the reduced effective mass is different in both cases. We can expand this equation for small wave vectors, keeping only terms up to k^6 , and find

$$E_4 = -\frac{\hbar^2 k^2}{2m_0 m_{lh}^*} \left[1 - \frac{\hbar^2 k^2}{2m_0} \frac{m_{lh}^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{m_{lh}^*}{4\mu_{lh}^3 E_0^2} \right]. \quad (\text{S125})$$

for the light hole band and

$$\begin{aligned} E_3 &= E_0 + \frac{\hbar^2 k^2}{2m_0 m_e^*} \times \\ &\quad \times \left[1 - \frac{\hbar^2 k^2}{2m_0} \frac{m_e^*}{4\mu_{lh}^2 E_0} + \left(\frac{\hbar^2 k^2}{2m_0} \right)^2 \frac{m_e^*}{4\mu_{lh}^3 E_0^2} \right] \quad (\text{S126}) \end{aligned}$$

for the electron band.

It is tempting to use these expressions, although they were derived as approximations for two special cases, also for the general case of band gaps and spin-orbit splittings, with the light hole, electron, and reduced effective masses obtained from the perturbative solution, as given by Eqs. (S110), (S118), and (S119). This provides an elegant method of interpolation between the small and large spin-orbit approximations. One might also use the

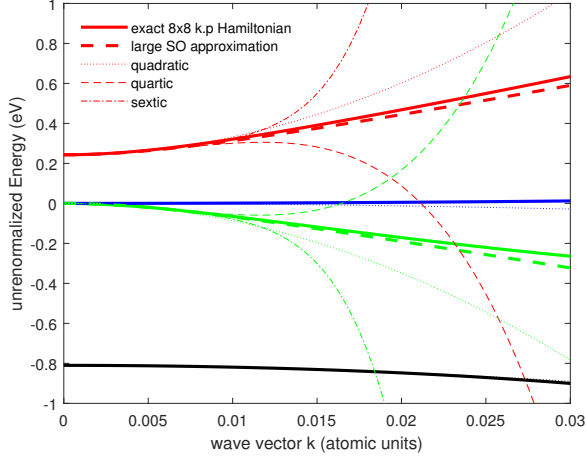


FIG. S9. Similar to Fig. S6 for InSb, but calculated with the reduced mass expressions (S124), (S125), and (S126). Thick lines show the heavy hole (blue), light hole (green), split-off hole (black), and electron bands (red) of the cubic characteristic equation given by Eq. (S95) (solid) and from the large spin-orbit approximation (S124) (dashed). Thin lines show the expansion of the square root including terms proportional to k^2 (dotted), k^4 (dashed), and k^6 (dot-dashed). Parabolic bands for the heavy and split off holes with experimental masses are also shown.

experimental masses or those calculated from a larger $\vec{k} \cdot \vec{p}$ Hamiltonian for improved accuracy.

Whether this approach will work, however, is not guaranteed and needs to be checked by comparison with more precise calculations. It is particularly problematic to use the electron (or light hole) mass from one type of approximation and the reduced mass from another.

Results are shown in Figs. S9 and S10 for InSb. Only the exact solutions (S95) to the cubic equation and the square-root expression (S124) in the approximation for large spin-orbit splittings yield reasonable results. The expansion of the square root up to terms linear in k^6 only results in a good approximation for a wave vector ka_B below 0.01 (energies of no more than 50 meV above the conduction band minimum). Due to the small band gap, nonparabolicity effects are very large. Because of the small gap, the expansion of the square root diverges very quickly and the coefficients become larger for higher-order terms.

S11. NONPARABOLICITY PARAMETERS

So far, our focus has been to find expressions for the band energies as a function of wave vector and we have plotted these results with various approximations derived from the 4×4 $\vec{k} \cdot \vec{p}$ Kane Hamiltonian (S57). Sometimes, it is also necessary to write the wave vector as a function

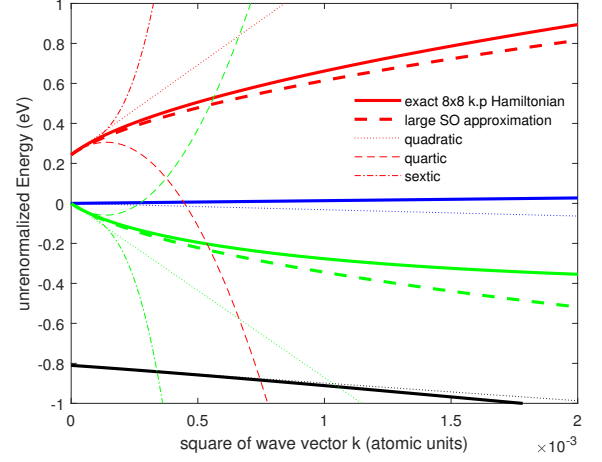


FIG. S10. Same as Fig. S9, but drawn with the square of the wave vector (in atomic units, i.e., inverse Bohr radii) on the horizontal axis.

of energy

$$\begin{aligned} \frac{\hbar^2 k^2}{2m_0 m_i} &= \frac{K^2}{m_i} = \epsilon_i (1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2) = \\ &= \epsilon_i \left(1 + \alpha'_i \frac{\epsilon_i}{E_0} + \beta'_i \frac{\epsilon_i^2}{E_0^2} \right), \end{aligned} \quad (\text{S127})$$

where i is the band index and the α s and β s are the lowest-order nonparabolicity coefficients. The units of the unprimed coefficients are an inverse power of energy, while the primed coefficients are dimensionless. For the heavy and light hole bands, we define $E_{\text{lh, hh}} = -\epsilon_{\text{lh, hh}}$ to keep the effective masses positive. For the electron and split-off hole bands, ϵ_i is the excess energy above or below the band extremum, i.e., $E_e = E_0 + \epsilon_e$ for the conduction band and $E_{\text{so}} = -\Delta_0 - \epsilon_{\text{so}}$ for the split-off hole band. To simplify the notation, we also define $K^2 = \hbar^2 k^2 / 2m_0$, which implies the use of atomic units (Bohr radius and Hartree as units for length and energy).

Since Vieta's solution (S95) to the cubic characteristic equation (S59) is "complicated" (see Bartoli 1983), we begin by deriving the nonparabolicity parameters in the small and large spin-orbit splitting approximations.

A. For small spin-orbit splittings

Conwell and Vassell (Phys. Rev. **166**, 797, 1968) introduce a non-parabolicity parameter α , which is required for the calculation of the joint density of states. Since they are interested primarily in GaAs, where $\Delta_0 \ll E_0$, they start with the characteristic equation for the conduction band for small spin-orbit splitting (S36)

$$\epsilon_e = K^2 + \frac{E_0}{2} \left(\sqrt{1 + 4K^2 \frac{E_P}{E_0^2}} - 1 \right), \quad (\text{S128})$$

where ϵ_e is the energy referenced to the conduction band minimum at the Γ -point. This can be solved as a function of the square of the wave vector as follows:

$$\left(\epsilon_e - K^2 + \frac{E_0}{2}\right)^2 = \frac{E_0^2}{4} \left(1 + 4K^2 \frac{E_P}{E_0^2}\right), \quad (\text{S129})$$

$$K^4 - K^2 E_0 - 2K^2 \epsilon_e + \epsilon_e E_0 + \epsilon_e^2 + \frac{1}{4} E_0^2 = \frac{1}{4} E_0^2 + K^2 E_P, \quad (\text{S130})$$

$$K^4 - K^2 (E_0 + E_P + 2\epsilon_e) + \epsilon_e (E_0 + \epsilon_e) = 0, \quad (\text{S131})$$

This biquadratic equation has the solution

$$\begin{aligned} K^2 &= \frac{1}{2} [(E_0 + E_P + 2\epsilon_e) \pm \\ &\pm \sqrt{(E_0 + E_P + 2\epsilon_e)^2 - 4\epsilon_e (E_0 + \epsilon_e)}] = \\ &= \epsilon_e + \frac{1}{2} \left[(E_0 + E_P) \pm \sqrt{(E_0 + E_P)^2 + 4\epsilon_e E_P} \right] = \\ &= \epsilon_e + \frac{E_0 + E_P}{2} \left[1 - \sqrt{1 + \frac{4\epsilon_e E_P}{(E_0 + E_P)^2}} \right]. \quad (\text{S132}) \end{aligned}$$

We have selected the solution with the minus sign, because we are looking for a solution where $\epsilon_e=0$ for $k=0$. This equation (S132) is the same as Eq. (2.2) by Conwell and Vassell (1968).

For sufficiently small energies ϵ_e , we can expand the square root to find

$$\begin{aligned} K^2 &= \epsilon_e - \frac{\epsilon_e E_P}{E_0 + E_P} + \frac{\epsilon_e^2 E_P^2}{(E_0 + E_P)^3} - \frac{2\epsilon_e^3 E_P^3}{(E_0 + E_P)^5} = \\ &= \epsilon_e \left[1 - \frac{E_P}{E_0 + E_P} + \frac{\epsilon_e E_P^2}{(E_0 + E_P)^3} - \frac{2\epsilon_e^2 E_P^3}{(E_0 + E_P)^5} \right] = \\ &= \epsilon_e \left[\frac{E_0}{E_0 + E_P} + \frac{\epsilon_e E_P^2}{(E_0 + E_P)^3} - 2 \frac{\epsilon_e^2 E_P^3}{(E_0 + E_P)^5} \right] = \\ &= \epsilon_e \left[m_e^* + \frac{\epsilon_e E_P^2}{(E_0 + E_P)^3} - 2 \frac{\epsilon_e^2 E_P^3}{(E_0 + E_P)^5} \right] = \\ &= m_e^* \epsilon_e \left[1 + \frac{\epsilon_e E_P^2}{m_e^* (E_0 + E_P)^3} - 2 \frac{\epsilon_e^2 E_P^3}{m_e^* (E_0 + E_P)^5} \right] = \\ &= m_e^* \epsilon_e \left[1 + \frac{\epsilon_e E_P^2}{E_0 (E_0 + E_P)^2} - 2 \frac{\epsilon_e^2 E_P^3}{E_0 (E_0 + E_P)^4} \right] \quad (\text{S133}) \end{aligned}$$

We have used Eq. (S40) for the effective electron mass within the small spin-orbit approximation. Equation (S133) is the same as Eq. (2.3) in Conwell and Vassell (1968).

We take the effective mass to the left hand side to arrive at the definition of the non-parabolicity parameters α_e and β_e for the conduction band

$$\frac{K^2}{m_e^*} = \epsilon_e \left[1 + \frac{\epsilon_e E_P^2}{E_0 (E_0 + E_P)^2} - 2 \frac{\epsilon_e^2 E_P^3}{E_0 (E_0 + E_P)^4} \right] =$$

$$\begin{aligned} &= \epsilon_e (1 + \alpha_e \epsilon_e + \beta_e \epsilon_e^2) = \\ &= \epsilon_e \left(1 + \alpha'_e \frac{\epsilon_e}{E_0} + \beta'_e \frac{\epsilon_e^2}{E_0^2} \right), \quad (\text{S134}) \end{aligned}$$

following the definition given by Eq. (2.4) in Conwell and Vassell (1968) or (for the primed quantities) the definition given by Bartoli (1983) in Eq. (A2). By comparing the coefficients, we find

$$\alpha_e = \frac{E_P^2}{E_0 (E_0 + E_P)^2} \approx \frac{1}{E_0}, \quad (\text{S135})$$

$$\alpha'_e = \frac{E_P^2}{(E_0 + E_P)^2} \approx 1, \quad (\text{S136})$$

$$\beta_e = -\frac{2E_P^3}{E_0 (E_0 + E_P)^4} \approx -\frac{2}{E_0 E_P}, \quad (\text{S137})$$

$$\beta'_e = -\frac{2E_0 E_P^3}{(E_0 + E_P)^4} \approx -\frac{2E_0}{E_P}. \quad (\text{S138})$$

These expressions were given by Conwell and Vassell (1968). The approximations are valid for $\epsilon_e \ll E_0$ and $E_0 \ll E_P$, which is usually the case. It seems that Eq. (S127) converges faster than Eq. (S45). Equations (S135-S138) agree with Eq. (A3a-A3b) in Bartoli (1983) for $\Delta_0=0$.

We can apply the same formalism to the light hole band, where we start with Eq. (S36)

$$\begin{aligned} E_{lh} &= K^2 + \frac{E_0}{2} \left(1 - \sqrt{1 + 4K^2 \frac{E_P}{E_0^2}} \right) = \\ &= K^2 - \frac{E_0}{2} \left(\sqrt{1 + 4K^2 \frac{E_P}{E_0^2}} - 1 \right). \quad (\text{S139}) \end{aligned}$$

This is the same equation as Eq. (S128) except that E_0 has the opposite sign. We also need to consider that the light hole energy E_{lh} is negative and its mass m_{lh} is positive. The outcome is therefore

$$\begin{aligned} K^2 &= E_{lh} \left[-\frac{E_0}{E_P - E_0} + \frac{E_{lh} E_P^2}{(E_P - E_0)^3} - 2 \frac{E_{lh}^2 E_P^3}{(E_P - E_0)^5} \right] = \\ &= -E_{lh} \left[\frac{E_0}{E_P - E_0} - \frac{E_{lh} E_P^2}{(E_P - E_0)^3} + 2 \frac{E_{lh}^2 E_P^3}{(E_P - E_0)^5} \right] = \\ &= -E_{lh} \left[m_{lh} - \frac{E_{lh} E_P^2}{(E_P - E_0)^3} + 2 \frac{E_{lh}^2 E_P^3}{(E_P - E_0)^5} \right]. \quad (\text{S140}) \end{aligned}$$

By bringing the light hole effective mass over to the left hand side and using Eq. (S42), we obtain

$$\begin{aligned} \frac{K^2}{m_{lh}^*} &= -E_{lh} \left[1 - \frac{E_{lh} E_P^2}{E_0 (E_P - E_0)^2} + 2 \frac{E_{lh}^2 E_P^3}{E_0 (E_P - E_0)^4} \right] = \\ &= -E_{lh} (1 - \alpha_{lh} E_{lh} + \beta_{lh} E_{lh}^2) = \\ &= -E_{lh} \left(1 - \alpha'_{lh} \frac{E_{lh}}{E_0} + \beta'_{lh} \frac{E_{lh}^2}{E_0^2} \right), \quad (\text{S141}) \end{aligned}$$

By comparing the coefficients, we find

$$\alpha_{lh} = \frac{E_P^2}{E_0(E_P - E_0)^2} \approx \frac{1}{E_0}, \quad (\text{S142})$$

$$\alpha'_{lh} = \frac{E_P^2}{(E_P - E_0)^2} \approx 1, \quad (\text{S143})$$

$$\beta_{lh} = \frac{2E_P^3}{E_0(E_P - E_0)^4} \approx \frac{2}{E_0E_P}, \quad (\text{S144})$$

$$\beta'_{lh} = \frac{2E_0E_P^3}{(E_P - E_0)^4} \approx \frac{2E_0}{E_P}. \quad (\text{S145})$$

Bartoli (1983) mentions that applying the small spin-orbit approximation to the light hole bands is not meaningful, since the light hole and split-off hole bands are strongly coupled for small spin orbit splittings. We have shown this already in Fig. S5.

B. For large spin-orbit splittings

Next, we discuss how the conduction band nonparabolicity parameters (S135-S138) need to be adjusted in the limit of large spin-orbit splittings. The characteristic equation (S61)

$$E_{3,4} = \frac{\hbar^2 k^2}{2m_0} + \frac{E_0}{2} \left(1 \pm \sqrt{1 + \frac{8}{3} \frac{\hbar^2 k^2 E_P}{2m_0 E_0^2}} \right) \quad (\text{S146})$$

has an additional factor $\frac{2}{3}$ under the square root. We rewrite this equation for the conduction band as

$$\epsilon_e = K^2 + \frac{E_0}{2} \left(\sqrt{1 + \frac{8}{3} K^2 \frac{E_P}{E_0^2}} - 1 \right) \quad (\text{S147})$$

to show the similarity with Eq. (S128) for vanishing spin-orbit splitting. We need to solve this equation for K^2 , just like in the case for vanishing spin-orbit splitting.

$$\left(\epsilon_e - K^2 + \frac{E_0}{2} \right)^2 = \frac{E_0^2}{4} \left(1 + \frac{8}{3} K^2 \frac{E_P}{E_0^2} \right), \quad (\text{S148})$$

$$K^4 - K^2 E_0 - 2K^2 \epsilon_e + \epsilon_e E_0 + \epsilon_e^2 + \frac{1}{4} E_0^2 = \frac{1}{4} E_0^2 + \frac{2}{3} K^2 E_P, \quad (\text{S149})$$

$$K^4 - K^2 \left(E_0 + \frac{2}{3} E_P + 2\epsilon_e \right) + \epsilon_e (E_0 + \epsilon_e) = 0, \quad (\text{S150})$$

$$K^2 = \frac{1}{2} \left[\left(E_0 + \frac{2}{3} E_P + 2\epsilon_e \right) \pm \sqrt{\left(E_0 + \frac{2}{3} E_P + 2\epsilon_e \right)^2 - 4\epsilon_e (E_0 + \epsilon_e)} \right] =$$

$$\begin{aligned} &= \epsilon_e + \frac{1}{2} \left[\left(E_0 + \frac{2}{3} E_P \right) \pm \right. \\ &\quad \left. \pm \sqrt{\left(E_0 + \frac{2}{3} E_P \right)^2 + \frac{8}{3} \epsilon_e E_P} \right] = \\ &= \epsilon_e + \frac{E_0 + \frac{2}{3} E_P}{2} \left[1 - \sqrt{1 + \frac{\frac{8}{3} \epsilon_e E_P}{\left(E_0 + \frac{2}{3} E_P \right)^2}} \right] \end{aligned} \quad (\text{S151})$$

For sufficiently small energies ϵ_e , we can expand the square root. The derivation works exactly the same way as for small spin-orbit splitting, except that E_P needs to be replaced by $\frac{2}{3} E_P$. We obtain

$$\alpha_e = \frac{\left(\frac{2}{3} E_P \right)^2}{E_0 \left(E_0 + \frac{2}{3} E_P \right)^2} \approx \frac{1}{E_0}, \quad (\text{S152})$$

$$\alpha'_e = \frac{\left(\frac{2}{3} E_P \right)^2}{\left(E_0 + \frac{2}{3} E_P \right)^2} \approx 1, \quad (\text{S153})$$

$$\beta_e = -\frac{2 \left(\frac{2}{3} E_P \right)^3}{E_0 \left(E_0 + \frac{2}{3} E_P \right)^4} \approx -\frac{3}{E_0 E_P}, \quad (\text{S154})$$

$$\beta'_e = -\frac{2E_0 \left(\frac{2}{3} E_P \right)^3}{\left(E_0 + \frac{2}{3} E_P \right)^4} \approx -\frac{3E_0}{E_P}. \quad (\text{S155})$$

We find that the first-order nonparabolicity coefficient α'_e is still about one, but the second-order coefficient β'_e has increased by a factor of $\frac{3}{2}$. We expect the same convergence criteria $\epsilon_e \ll E_0$ and $E_0 \ll E_P$ as for small spin-orbit splittings.

We can use the same approach for the light hole band and find for large spin-orbit splittings

$$\alpha_{lh} = \frac{\left(\frac{2}{3} E_P \right)^2}{E_0 \left(\frac{2}{3} E_P - E_0 \right)^2} \approx \frac{1}{E_0}, \quad (\text{S156})$$

$$\alpha'_{lh} = \frac{\left(\frac{2}{3} E_P \right)^2}{\left(\frac{2}{3} E_P - E_0 \right)^2} \approx 1, \quad (\text{S157})$$

$$\beta_{lh} = \frac{2 \left(\frac{2}{3} E_P \right)^3}{E_0 \left(\frac{2}{3} E_P - E_0 \right)^4} \approx \frac{3}{E_0 E_P}, \quad (\text{S158})$$

$$\beta'_{lh} = \frac{2E_0 \left(\frac{2}{3} E_P \right)^3}{\left(\frac{2}{3} E_P - E_0 \right)^4} \approx \frac{3E_0}{E_P}. \quad (\text{S159})$$

These equations agree with Bartoli (1983), Eqs. (A3a-A6), in the limit of large spin-orbit splittings. For InSb, this expansion should produce better results than for GaAs, because the interaction of the light and split-off hole bands is weaker in InSb, see Fig. S6.

C. Using the electron/light hole reduced mass

For this section, it is important to use the masses obtained from the large spin-orbit splitting approximation

consistently. If we mix and match the masses from different approximations, then Eq. (S163) will not be valid.

We can also start the calculation of the nonparabolicity parameters α and β with Eqs. (S44) and (S69) for the conduction band, which read

$$\epsilon_e = K^2 + \frac{E_0}{2} \left(\sqrt{1 + K^2 \frac{2}{\mu_{lh} E_0}} - 1 \right), \quad (\text{S160})$$

where $\epsilon_e = E - E_0$ is again the energy of the electron above the conduction band minimum. This leads us to the biquadratic equation

$$K^4 - K^2 \left(E_0 + \frac{E_0}{2\mu_{lh}} + 2\epsilon_e \right) + \epsilon_e (\epsilon_e + E_0) = 0, \quad (\text{S161})$$

which has the solution

$$K^2 = \epsilon_e + \frac{E_0}{2} \left(1 + \frac{1}{2\mu_{lh}} \right) \left[1 - \sqrt{1 + \frac{2\epsilon_e}{\mu_{lh} E_0 \left(1 + \frac{1}{2\mu_{lh}} \right)^2}} \right]. \quad (\text{S162})$$

We recognize the term in round parentheses as the inverse effective electron mass given by Eqs. (S51) and (S77). This simplifies the expression for K^2 to read

$$K^2 = \epsilon_e + \frac{E_0}{2m_e^*} \left(1 - \sqrt{1 + \frac{2\epsilon_e m_e^{*2}}{\mu_{lh} E_0}} \right). \quad (\text{S163})$$

By expanding the square root in orders of ϵ_e , we obtain

$$K^2 = \epsilon_e \left(1 - \frac{m_e^*}{2\mu_{lh}} + \frac{m_e^{*3} \epsilon_e}{4\mu_{lh}^2 E_0} - \frac{m_e^{*5} \epsilon_e^2}{4\mu_{lh}^3 E_0^2} \right). \quad (\text{S164})$$

Equations (S51) and (S77) imply that the first two terms inside the parentheses are equal to the effective electron mass:

$$m_e^* = 1 - \frac{m_e^*}{2\mu_{lh}}. \quad (\text{S165})$$

The wave vector as a function of energy can therefore be obtained from

$$K^2 = m_e^* \epsilon_e \left(1 + \frac{m_e^{*2} \epsilon_e}{4\mu_{lh}^2 E_0} - \frac{m_e^{*4} \epsilon_e^2}{4\mu_{lh}^3 E_0^2} \right). \quad (\text{S166})$$

The nonparabolicity parameters for the electron band are therefore given by, with reference to Eq. (S165),

$$\alpha_e = \frac{m_e^{*2}}{4\mu_{lh}^2 E_0} = \frac{(1 - m_e^*)^2}{E_0}, \quad (\text{S167})$$

$$\alpha'_e = \frac{m_e^{*2}}{4\mu_{lh}^2} = \left(\frac{m_e^*}{2\mu_{lh}} \right)^2 = (1 - m_e^*)^2, \quad (\text{S168})$$

$$\beta_e = -\frac{m_e^{*4}}{4\mu_{lh}^3 E_0^2} = \frac{-2m_e^* (1 - m_e^*)^3}{E_0^2}, \quad (\text{S169})$$

$$\beta'_e = -\frac{m_e^{*4}}{4\mu_{lh}^3} = -2m_e^* (1 - m_e^*)^3. \quad (\text{S170})$$

These nonparabolicity parameters were derived for both the small and large spin-orbit splitting approximations. It is tempting to use them also for the case of general spin-orbit splittings, since they offer an elegant way of interpolating between both limits. We note that Eqs. (S168) and (S170) are the leading factors in the nonparabolicity expressions given by Bartoli (1983) in the limit of large spin-orbit splittings. Compare also Eqs. (1f) and (6a) in Palik, Picus, Teitler, and Wallis (1961). α'_e is approximately equal to unity, since the effective electron mass is small. β'_e is on the order of the effective electron mass, i.e., much smaller than α'_e .

For the light hole band, the energy is

$$E_{lh} = K^2 - \frac{E_0}{2} \left(\sqrt{1 + K^2 \frac{2}{\mu_{lh} E_0}} - 1 \right), \quad (\text{S171})$$

which leads to the biquadratic equation

$$K^4 - K^2 \left(-E_0 + \frac{E_0}{2\mu_{lh}} + 2E_{lh} \right) + E_{lh} (E_{lh} - E_0) = 0. \quad (\text{S172})$$

This equation is similar to Eq. (S161) for the conduction band, except that both E_0 and μ_{lh} appear with the opposite sign. The solution is therefore

$$K^2 = E_{lh} - \frac{E_0}{2} \left(1 - \frac{1}{2\mu_{lh}} \right) \left[1 - \sqrt{1 + \frac{2E_{lh}}{\mu_{lh} E_0 \left(1 - \frac{1}{2\mu_{lh}} \right)^2}} \right]. \quad (\text{S173})$$

The term in round parentheses is related to the inverse light hole effective mass through Eqs. (S48) and (S74). This simplifies the expression for K^2 to read

$$K^2 = E_{lh} + \frac{E_0}{2m_{lh}^*} \left[1 - \sqrt{1 + \frac{2E_{lh} m_{lh}^{*2}}{\mu_{lh} E_0}} \right]. \quad (\text{S174})$$

By expanding the square root in orders of E_{lh} , we obtain

$$K^2 = E_{lh} \left(1 - \frac{m_{lh}^*}{2\mu_{lh}} + \frac{m_{lh}^{*3} E_{lh}}{4\mu_{lh}^2 E_0} - \frac{m_{lh}^{*5} E_{lh}^2}{4\mu_{lh}^3 E_0^2} \right). \quad (\text{S175})$$

Equations (S48) and (S74) imply that the first two terms inside the parentheses are related to the effective light hole mass by

$$-m_{lh}^* = 1 - \frac{m_{lh}^*}{2\mu_{lh}}. \quad (\text{S176})$$

If we remember the definition $\epsilon_{lh} = -E_{lh}$ from the beginning of Sec. S11, then the wave vector as a function of energy can be obtained from

$$K^2 = m_{lh}^* \epsilon_{lh} \left(1 + \frac{m_{lh}^{*2} \epsilon_{lh}}{4\mu_{lh}^2 E_0} + \frac{m_{lh}^{*4} \epsilon_{lh}^2}{4\mu_{lh}^3 E_0^2} \right). \quad (\text{S177})$$

Equations (S166) and (S177) are not identical, if we change the band subscripts for the energy and for the

TABLE SI. Effective masses and nonparabolicity parameters for InSb at low temperature, calculated with the values $E_0=0.237$ eV, $\Delta_0=0.81$ eV, and $E_P=23.1$ eV taken from Lawaetz (1971). The top row uses the limit of large spin-orbit splittings in Sec. S7 with Eqs. (S65,S66,S68,S168,S170,S179,S181). The second row uses the perturbative approximation in Sec. S9 with Eqs. (S110,S118,S119,S168,S170,S179,S181). The third row lists results with finite spin-orbit splittings from the Appendix in Bartoli (1983).

m_e^*	m_{lh}^*	μ_{lh}	α'_e	β'_e	α'_{lh}	β'_{lh}
0.0152	0.0156	0.00769	0.970	-0.029	1.032	0.033
0.0136	0.0156	0.00728	0.973	-0.026	1.032	0.033
0.0136	0.0156	0.00728	0.896	-0.075	1.182	0.264

effective mass (but not for the reduced mass), because they have the opposite sign for the quartic nonparabolicity parameter β' .

The nonparabolicity parameters for the light hole band are therefore given by, with reference to Eq. (S176),

$$\alpha_{lh} = \frac{m_{lh}^{*2}}{4\mu_{lh}^2 E_0} = \frac{(1 + m_{lh})^2}{E_0}, \quad (\text{S178})$$

$$\alpha'_{lh} = \frac{m_{lh}^{*2}}{4\mu_{lh}^2} = \left(\frac{m_{lh}^*}{2\mu_{lh}}\right)^2 = (1 + m_{lh}^*)^2, \quad (\text{S179})$$

$$\beta_{lh} = \frac{m_{lh}^{*4}}{4\mu_{lh}^3 E_0^2} = \frac{2m_{lh}^* (1 + m_{lh}^*)^3}{E_0^2}, \quad (\text{S180})$$

$$\beta'_{lh} = \frac{m_{lh}^{*4}}{4\mu_{lh}^3} = 2m_{lh}^* (1 + m_{lh}^*)^3. \quad (\text{S181})$$

Equations. (S179) and (S181) are the leading factors in the nonparabolicity expressions given by Bartoli (1983) in the limit of large spin-orbit splittings.

The expressions given in this section are evaluated as shown in Fig. S11 for the electron and light hole bands of InSb. The parabolic approximation (dotted) is poor except for very small energies above the conduction band minimum or below the valence band maximum. Keeping only the lowest-order nonparabolic term α_i is a good approximation to the exact expression containing the square root. If the β_i -term is included (dot-dashed), the deviation from the exact square-root expression (S163) is nearly indistinguishable on this scale. This energy dispersion is quite similar to that calculated by Kane (1957), Fig. 1, using an $8 \times 8 \vec{k} \cdot \vec{p}$ -model with corrections due to interactions with higher-energy bands.

While the light hole mass of $m_{lh}=0.0156$ given by the large spin-orbit splitting approximation is similar to the experimental mass, the electron mass of $m_e=0.0152$ is about 10% too high. But inserting the experimental electron mass ($m_e=0.013$ or 0.014) in Eq. (S163) yields inconsistent results, because different approximations are mixed in the expressions, with invalid outcomes.

Table SI shows the effective electron and light hole masses and nonparabolicity parameters calculated in dif-

ferent approximations.

D. Bartoli's expressions

Without much discussion and no derivation or references, Bartoli *et al.* (1983) give the following expressions for the effective masses and the non-parabolicity:

$$\delta = \frac{\Delta_0}{E_0}, \quad (\text{S182})$$

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{3} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right), \quad (\text{S183})$$

$$m_{lh}^* = \frac{(\delta + 1.5) m_e^*}{\delta + 1 - m_e^* (2\delta + 2.5)}, \quad (\text{S184})$$

$$\alpha_e = \frac{(1 - m_e^*)^2 (1.5 + 2\delta + \delta^2)}{(1.5 + \delta)(1 + \delta)}, \quad (\text{S185})$$

$$\beta_e = \frac{-2(1 - m_e^*)^3}{(1.5 + \delta)^2 (1 + \delta)} \left[\frac{\delta^2}{4} + \frac{m_e^* (1.5 + 2\delta + \delta^2)^2}{1 + \delta} \right] \quad (\text{S186})$$

$$\alpha_{lh} = (1 + m_{lh}^*)^2 \left(1 + \frac{1}{2\delta} \right), \quad (\text{S187})$$

$$\beta_{lh} = (1 + m_{lh}^*)^3 \left[\frac{1}{2\delta} + \frac{3}{4\delta^2} + 2m_{lh}^* \left(1 + \frac{1}{2\delta} \right)^2 \right] \quad (\text{S188})$$

They mention that Eq. (S185) was previously given by Palik *et al.* (1961), but that is not immediately easy to see.

The square of the wave vector as function of the excess energy ϵ_i evaluated using the Bartoli expressions is shown in Fig. S12. The splitting between the electron and light hole bands is larger for these expressions than for the large spin-orbit splitting approximation, because the second-order corrections involving the β_i -terms are larger. Overall, the agreement with the bands calculated by Kane (1957) is worse than for the pure approximation for large spin-orbit splittings. Therefore, it does not seem beneficial to use the Bartoli expressions for InSb.

E. Expression given by Menendez (2020)

Menendez (2020) introduced a nonparabolicity parameter for the conduction band

$$\Delta_\Gamma = \frac{3}{2} \left(\frac{2}{E_0} + \frac{1}{E_0 + \Delta_0} \right)^{-1}. \quad (\text{S189})$$

In our notation, $2\Delta_\Gamma = 1/\alpha_e$. Therefore, Δ_Γ should approach $E_0/2$ for both small and large spin-orbit splittings. This condition is fulfilled for small spin-orbit splittings, but not for large spin-orbit splittings, where $\Delta_\Gamma = 3E_0/4$. Therefore, the use of Δ_Γ for InSb is not recommended, because it will underestimate the non-parabolicity by up to 30%.

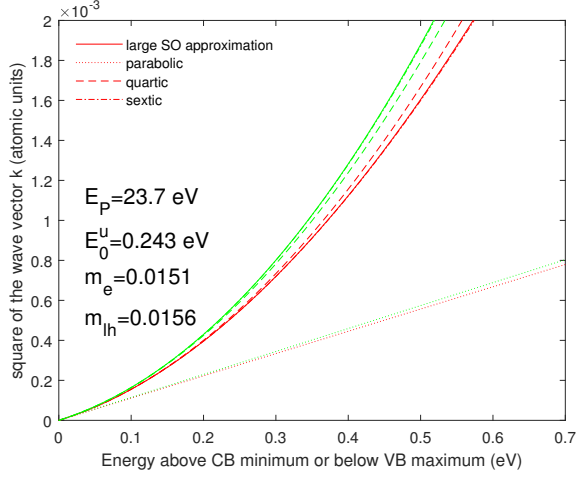


FIG. S11. The square of the wave vector (in atomic units) versus energy from the extremum for electrons (red) and light holes (green) in InSb, calculated using Eq. (S163) within the large spin-orbit splitting approximation (solid), using masses from Table SI. The dotted lines show the parabolic band dispersion. The dashed and dot-dashed lines show the next two terms in the Taylor expansions of the square root. Only the first term ($\beta_i=0$) gives a good approximation (dashed). If the β_i -term is included (dot-dashed), the deviation from the exact square-root expression (S163) is nearly indistinguishable on this scale.

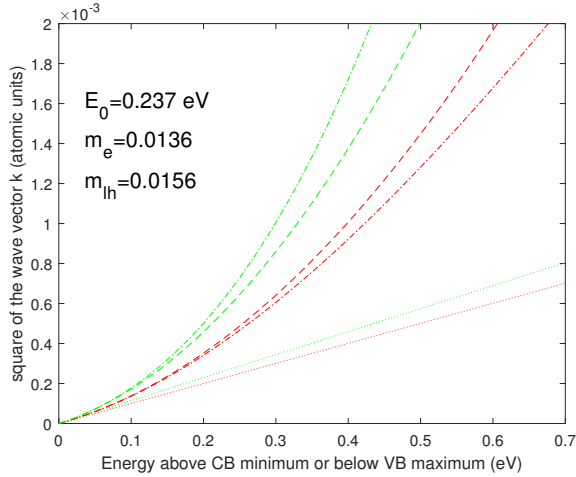


FIG. S12. As Fig. S11, but evaluated using the expressions given by Bartoli *et al.* (1983).

S12. DENSITY OF STATES

For the calculation of the chemical potential, we need the density of states

$$g_i(\epsilon_i) = \frac{1}{4\pi^3} \int d^3\vec{k} \delta(E_{i\vec{k}} - \epsilon_i) = \frac{1}{\pi^2} \int_0^\infty k^2 dk \delta(E_{ik} - \epsilon_i), \quad (\text{S190})$$

where i is the band index. We have included the spin degeneracy and assume that the bands are spherically symmetric.

By taking the derivative of Eq. (S127) on both sides, we find

$$dk = \sqrt{\frac{m_0 m_i}{2\hbar^2 \epsilon_i}} \frac{1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2}{\sqrt{1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2}} d\epsilon_i, \quad (\text{S191})$$

compare Bartoli (A7), and

$$k^2 dk = \frac{1}{2} \left(\frac{2m_0 m_i}{\hbar^2} \right)^{\frac{3}{2}} \times \sqrt{\epsilon_i (1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2)} (1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2) d\epsilon_i. \quad (\text{S192})$$

The density of states is therefore

$$g_i(\epsilon_i) = \frac{1}{2\pi^2} \left(\frac{2m_0 m_i}{\hbar^2} \right)^{\frac{3}{2}} \times \sqrt{\epsilon_i (1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2)} (1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2), \quad (\text{S193})$$

compare Fox (3.24) and Masut (2022), Eq. (6). We see that the nonparabolicity enhances the density of states by a factor

$$(1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2) \sqrt{1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2} \approx 1 + \frac{5}{2} \alpha_i \epsilon_i \quad (\text{S194})$$

to first order in $\alpha_i \epsilon_i$ if we set β_i to zero. Since the density of states depends on $m_i^{\frac{3}{2}}$, we can define an energy-dependent density-of-states effective mass

$$m_{i,\text{DOS}}(\epsilon_i) = m_i \sqrt[3]{1 + \alpha_i \epsilon_i + \beta_i \epsilon_i^2} (1 + 2\alpha_i \epsilon_i + 3\beta_i \epsilon_i^2)^{\frac{2}{3}}. \quad (\text{S195})$$

By setting $\beta_e=0$ and keeping only terms linear in $\alpha_e \epsilon_e$, the effective electron mass m_e increases approximately like

$$m_{e,\text{DOS}}(\epsilon) \approx m_e \left(1 + \frac{5}{3} \alpha_e \epsilon_e \right). \quad (\text{S196})$$

In other words, when the excess energy ϵ_e is equal to the band gap E_0 (i.e., $\alpha_e \epsilon_e \approx 1$), the effective electron mass m_e has nearly tripled. This is shown in Fig. S13, which plots the effective density-of-states electron mass of InSb as a function of excess energy above the conduction band minimum. Most of the mass enhancement is due to α_e -term (shown by the dotted line). We therefore have confidence that the expansion (S193) converges well in the approximation for large spin-orbit splittings. The energy dependence of the effective density-of-states light hole mass is also shown in Fig. S13.

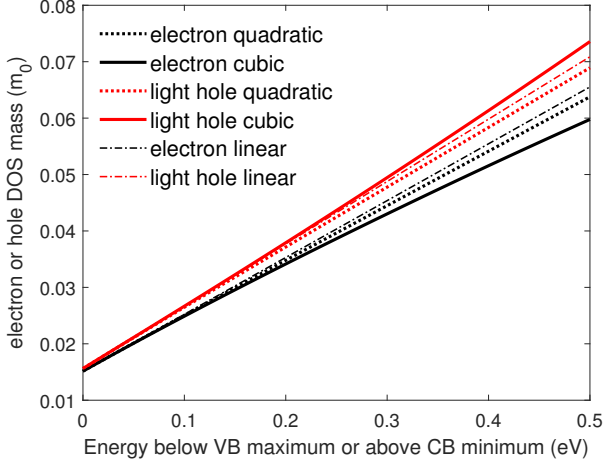


FIG. S13. Effective density of states electron mass m_e (black) and light hole mass m_{lh} (red) of InSb at 0 K as a function of excess energy above the conduction band minimum, calculated from Eq. (S195) (solid). The dotted lines show the results with β_e or β_{lh} set to zero. The dash-dotted lines show the linear expansion (S196).

S13. CHEMICAL POTENTIAL WITH NONPARABOLICITY

A. General expressions for all semiconductors

We apply the density of states (S193) for nonparabolic bands to calculate the chemical potential μ and the intrinsic carrier concentration n for InSb as a function of temperature T .

The electron density n at temperature T is given by

$$n(T) = \int_0^{\infty} d\epsilon g_e(\epsilon) f(E_0 + \epsilon_e), \quad (\text{S197})$$

where $f(E)$ is the Fermi-Dirac distribution function

$$f(E) = \frac{1}{\exp\left(\frac{E-\mu}{k_B T}\right) + 1}. \quad (\text{S198})$$

By setting $\beta_e=0$ and keeping only terms linear in α_e , we find that the density of states enhancement factor (S194) is approximately $1 + \frac{5}{2}\alpha_e\epsilon_e$. With the substitutions $y = \epsilon_e/k_B T$ and $x = (\mu - E_0)/k_B T$, the electron density can be written using Fermi-Dirac integrals as (compare Menendez 2020)

$$n(T) = N_e(T) \left[F_{\frac{1}{2}}\left(\frac{\mu - E_0}{k_B T}\right) + \frac{15}{4}\alpha_e k_B T F_{\frac{3}{2}}\left(\frac{\mu - E_0}{k_B T}\right) \right], \quad (\text{S199})$$

where the prefactor N_e is given by

$$N_e(T) = \frac{1}{4} \left(\frac{2m_0 m_e k_B T}{\pi \hbar^2} \right)^{3/2}. \quad (\text{S200})$$

Similarly, the light and heavy hole densities are given by

$$p_{lh}(T) = N_{lh}(T) \left[F_{\frac{1}{2}}\left(-\frac{\mu}{k_B T}\right) + \frac{15}{4}\alpha_{lh} k_B T F_{\frac{3}{2}}\left(-\frac{\mu}{k_B T}\right) \right],$$

$$p_{hh}(T) = N_{hh}(T) F_{\frac{1}{2}}\left(-\frac{\mu}{k_B T}\right), \quad (\text{S201})$$

if we ignore the nonparabolicity of the heavy hole band, with prefactors given by

$$N_{lh}(T) = \frac{1}{4} \left(\frac{2m_0 m_{lh} k_B T}{\pi \hbar^2} \right)^{3/2} \quad \text{and}$$

$$N_{hh}(T) = \frac{1}{4} \left(\frac{2m_0 m_{hh} k_B T}{\pi \hbar^2} \right)^{3/2}. \quad (\text{S202})$$

We can find the chemical potential by solving the equation

$$n(T) = p_{lh}(T) + p_{hh}(T) \quad (\text{S203})$$

at a given temperature, for example using polylogarithm functions in MATLAB. As an example, we show the electron and hole density of InSb at 300 K as a function of chemical potential in Fig. S14. For the holes, the light hole density is only a very small contribution, because the heavy hole is about 30 times heavier than the light hole. Therefore, the nonparabolicity correction for the hole bands does not matter much. The nonparabolicity correction for the electron concentration is sizeable, which can be seen from the difference between the red dotted and solid lines.

B. Evaluation of several scenarios for InSb

At 300 K with a band gap of 0.2 eV, $E_P=23.1$ eV, a heavy hole mass of 0.43, and electron and light hole masses calculated within the large spin-orbit splitting approximation, electron and hole concentrations become equal at a chemical potential of 161 meV (166 meV) with (without) the lowest order nonparabolic correction and α_e and α_{lh} from Table SI. Compare Rivero Arias (2023).

Two trends are seen in Fig. S14: First, the chemical potential is lower when the nonparabolicity is taken into account. Second, the carrier density (at this lower chemical potential) is higher than in the parabolic approximation. Both trends can be understood from the nondegenerate (classical Maxwell-Boltzmann) limit for parabolic bands (Sze 1981, Ashcroft & Mermin 1976)

$$\mu = \frac{E_0}{2} + \frac{3kT}{4} \ln\left(\frac{m_{dv}}{m_e}\right), \quad (\text{S204})$$

$$n = p = 2 \left(\frac{m_0 k_B T}{2\pi \hbar^2} \right)^{3/2} (m_e m_{dv})^{3/4} \exp\left(-\frac{E_0}{2k_B T}\right) \quad (\text{S205})$$

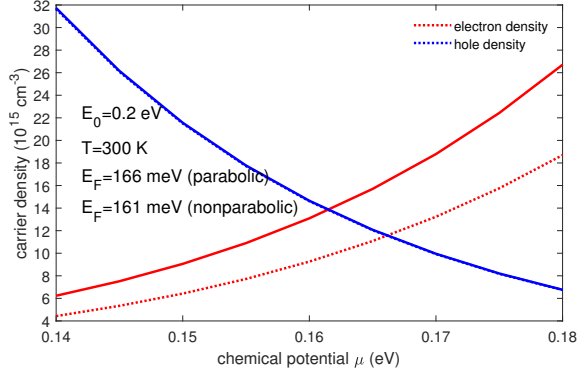


FIG. S14. Electron (red) and hole density (blue) as a function of chemical potential at 300 K in the parabolic approximation (dotted) and with the lowest nonparabolic corrections (solid). The band gap and heavy hole mass were assumed to be 0.2 eV and 0.43, respectively, and the electron and light hole mass were calculated in the large spin-orbit approximation.

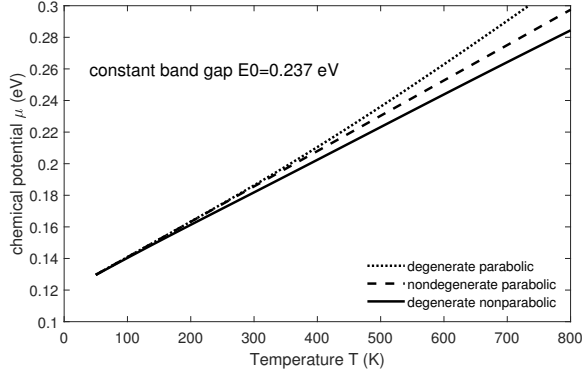


FIG. S15. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. A band gap of 0.237 eV and temperature-independent effective masses were assumed.

If the electron mass increases due to nonparabolicity while the density-of-states hole mass m_{dv} remains nearly constant, then the chemical potential decreases following Eq. (S204). At the same time, the carrier concentration increases like $m_e^{3/4}$ due to Eq. (S205).

Figures S15 and S16 show the chemical potential and the intrinsic carrier concentration versus temperature, assuming a band gap and effective masses that are independent of temperature. We again see that the chemical potential is lowered and the carrier concentration increased by nonparabolicity.

In our next model for the temperature dependence of the chemical potential and the intrinsic carrier concentration, we include the full temperature dependence of the band gap (Debye-Waller, self energy, and thermal expansion

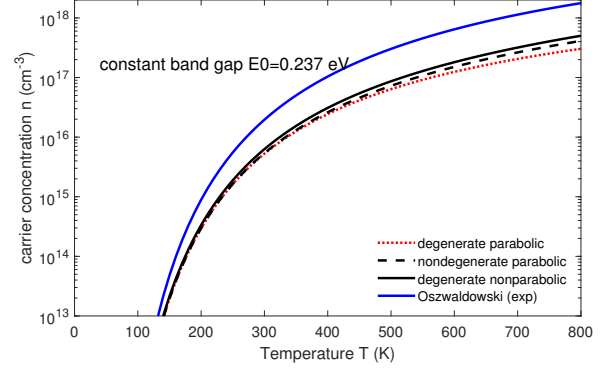


FIG. S16. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. A band gap of 0.237 eV and temperature-independent effective masses were assumed.

contributions) using the Bose-Einstein expression

$$E_0(T) = E_B - a_B \left[1 + \frac{2}{\exp(\Omega/k_B T) - 1} \right] \quad (\text{S206})$$

with parameters $E_B=261$ meV, $a_B=26$ meV, and $\Omega=18.9$ meV (correcting an error in the article by Rivero Arias, 2023). We ignore the small temperature dependence of the coupling parameter E_P due to thermal expansion, see Eq. (21) in Emminger (2022). The lower band gap at high temperatures will also reduce the light hole and electron masses according to Eq. (S31) and (S110). The temperature dependence of the light hole mass is not important, because the hole density is dominated by heavy holes. The effective mass of heavy holes should not change much with temperature, because this mass is dominated by the coupling with higher-energy conduction bands (see Dresselhaus, Kip, and Kittel, for example). The higher gaps only change by a small fraction. The temperature dependence of the heavy hole mass (both theory and experiment) for InSb has been discussed by Oswaldowski and Zimpel (1988). For our purposes, we select a constant value of the heavy hole mass (independent of temperature) and parabolic heavy hole bands.

The choice of the effective masses has been discussed in the literature. Some authors, for example Oswaldowski and Zimpel (1988), have argued that the unrenormalized band gap E_B should be used to calculate the effective masses within $\vec{k} \cdot \vec{p}$ -theory. At the moment, however, we take the view that the experimental band gap E_0 (sometimes also called the thermal gap or optical gap) should influence the effective light hole and electron masses, with a nearly constant matrix element E_P . The effective masses therefore decrease significantly with increasing temperature as the band gap closes according to Eq. (S206).

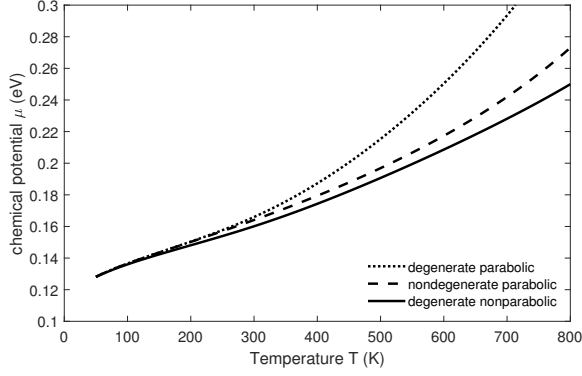


FIG. S17. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The full temperature dependence of the direct gap according to Eq. (S206) and of the light hole and electron masses was included.

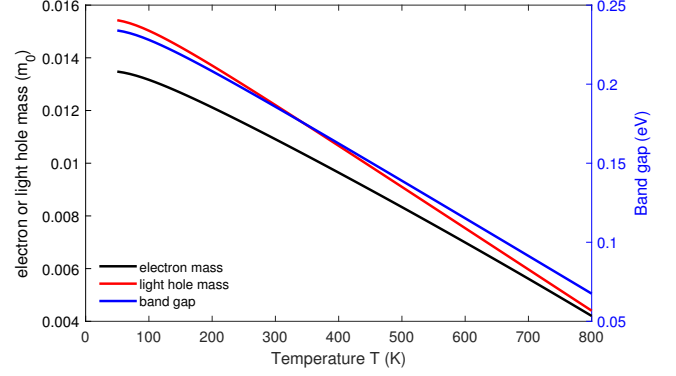


FIG. S19. Effective masses of the electron (black) and light hole (red) bands and band gap (blue) of InSb as a function of temperature.

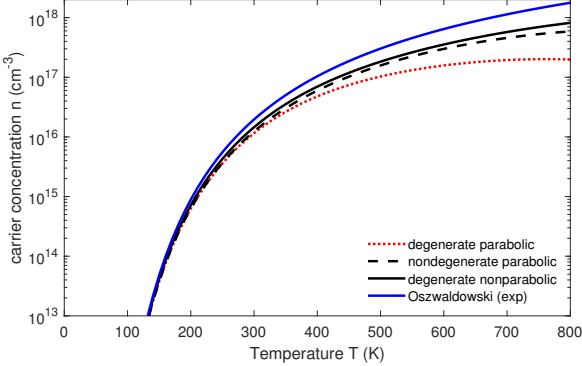


FIG. S18. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The full temperature dependence of the direct gap according to Eq. (S206) and of the light hole and electron masses was included. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

As shown in Fig. S17, the chemical potential increases superlinearly at high temperatures, where the band gap shrinks significantly, see Eq. (S206) and Fig. 3 in Rivero Arias (2023). The band gap reduction also decreases the effective electron mass, see Eq. (S31), and therefore increases the chemical potential according to Eq. (S204). The temperature dependence of the intrinsic carrier concentration is shown in Fig. S18 with the temperature dependence of the band gap and the effective masses considered. At 800 K, the carrier concentration is increased by about 33% compared to the calculation with a temperature-independent band gap and electron mass shown in Fig. S16, from 6×10^{17} to $8 \times 10^{17} \text{ cm}^{-3}$.

Oszwaldowski and Zimpel (1988) obtained the temperature dependence of the intrinsic carrier concentration of InSb from 200 to 800 K from Hall measurements, also shown in Fig. S18. Assuming a Hall factor of unity, they found an intrinsic carrier concentration near $1.8 \times 10^{18} \text{ cm}^{-3}$ at 800 K, about two times higher than the carrier concentration of $8.2 \times 10^{17} \text{ cm}^{-3}$ calculated with our degenerate nonparabolic model (with temperature-dependent band gap and effective masses). Only a much higher (perhaps temperature-independent) density of states mass for the electrons would explain these experimental data. The next order nonparabolicity correction would not improve the agreement, because $\beta_e < 0$.

For completeness, we show the temperature dependence of the band gap and of the effective electron and light hole masses in Fig. S19. As expected, all have a similar temperature dependence. It is interesting that the chemical potential at 800 K is about five times as large as the band gap.

We also calculated the temperature dependence of the chemical potential and the intrinsic carrier concentration of InSb by assuming that the thermal or optical gap (which enters the Fermi-Dirac distribution) varies with temperature, while the effective masses and nonparabolicity parameters are calculated using $\vec{k} \cdot \vec{p}$ -theory from the low-temperature band gap. The results are shown in Figs. S20 and S21. With this model, the calculated carrier density at 800 K is only 20% lower than the experimental value measured by Oszwaldowski and Zimpel (1988). It is also a remarkable coincidence that the degenerate nonparabolic carrier concentration is nearly the same as the nondegenerate parabolic result.

Next, we will calculate the chemical potential and the intrinsic carrier concentration using the thermal (or optical) gap from Eq. (S206) in the Fermi-Dirac distribution function, while considering only the thermal expansion contribution to the band gap changes in the calculation of the effective masses and the nonparabolicity param-

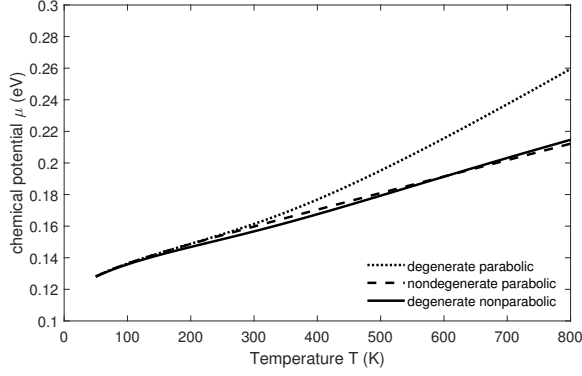


FIG. S20. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The full temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses and nonparabolicity parameters were kept constant at values calculated from the low-temperature band gap.

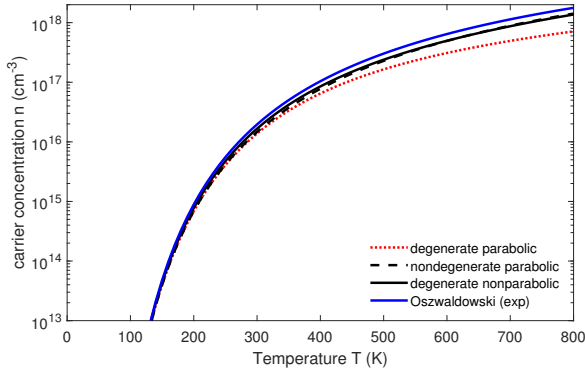


FIG. S21. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses and nonparabolicity parameters were kept constant at values calculated from the low-temperature band gap. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

ters, but not the redshift due to electron-phonon renormalization. The redshift of the direct gap due to thermal expansion is given by (Zollner 1991 GaSb)

$$\left(\frac{\partial E_0}{\partial T}\right)_{\text{TE}} = -3\alpha(T)B\left(\frac{\partial E_0}{\partial p}\right)_T, \quad (\text{S207})$$

where $\alpha(T)$ is the temperature-dependent thermal expansion coefficient taken from S. I. Novikova, *Thermal expansion*, in *Semiconductors and Semimetals*, vol. 2,

edited by R. K. Willardson and A. C. Beer, pp. 33-38 (1966), $B=46$ GPa the bulk modulus (assumed to be independent of temperature, LaBoe Vol. 17a, page 317), and the last factor equal to 0.155 eV/GPa is the pressure coefficient of the InSb band gap (Booth 1982, LaBoe Vol. 22a, page 124). For $\alpha(T)$ above room temperature, see Cai J. Appl. Phys. 2013.

The temperature dependence of the band gap due to thermal expansion and ignoring the Debye-Waller and self-energy contribution is therefore

$$E_0^{\text{TE}}(T) = E_0(T=0) - 3B\left(\frac{\partial E_0}{\partial p}\right)_T \int_0^T \alpha(\theta) d\theta. \quad (\text{S208})$$

We also call this the "mass band gap".

Following Roucka (PRB 81, 245214, 2010), we write the thermal expansion coefficient of zinc blende semiconductors as

$$\alpha(T) = A\left(\frac{T}{\Theta_D}\right)^3 I_D\left(\frac{\Theta_D}{T}\right), \quad (\text{S209})$$

where

$$I_D(x_D) = \int_0^{x_D} \frac{x^4 e^x dx}{(e^x - 1)^2} \quad (\text{S210})$$

is the Debye integral (which can be solved numerically in MATLAB), A is an adjustable parameter, and Θ_D the Debye temperature. The Debye temperature for InSb is about 205 K for InSb (LaBo 22a 1987, page 129), but we treat it as an adjustable parameter to fit the thermal expansion coefficients from Novikova (1966). With parameters $A=17.5 \times 10^{-6} \text{ K}^{-1}$ and $\Theta_D=450$ K, satisfactory agreement with the experimental data can be achieved, see Fig. S22.

An even better model for the temperature dependence of the thermal expansion coefficient, which also explains its low-temperature behavior, is given in the main text.

The temperature dependence of the band gap due to thermal expansion, calculated from Eq. (S208), is shown in Fig. S23. This is the same result as shown by Zollner (SSC, 1991).

The chemical potential and intrinsic carrier concentration of InSb as a function of temperature, taking into account only the thermal expansion contribution to the band gap shift (not the electron-phonon renormalization) when calculating the effective masses as a function of temperature, are shown in Figs. S24 and S25. Figure S26 shows the effective light hole and electron masses of InSb, the experimental band gap, and the thermal expansion band gap used to calculate the effective masses. As expected, the temperature dependence of the masses follows the dependence of the band gap due to thermal expansion in this situation. As shown by Fig. S25, the intrinsic carrier concentration is about 10^{18} cm^{-3} , lower than the result when keeping the masses constant as in Fig. S21 and higher than when calculating the temperature dependence of the masses from the experimental band gap as shown in Fig. S18.

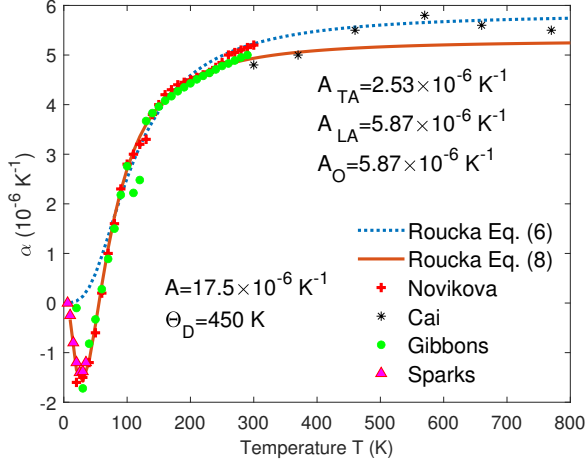


FIG. S22. Linear thermal expansion coefficient α versus temperature taken from Novikova (1966) (symbols) along with a fit to the data using Eq. (S209) (solid).

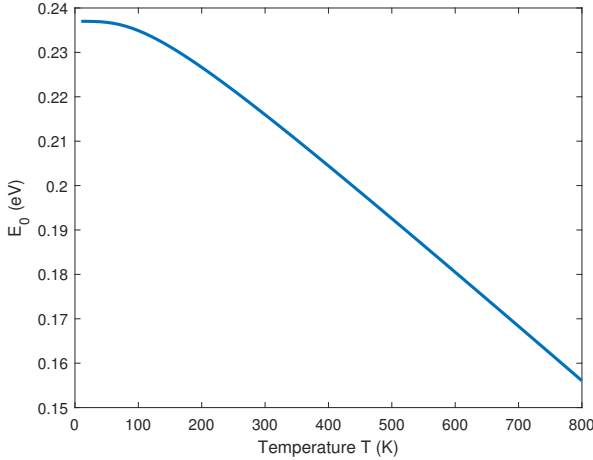


FIG. S23. Temperature dependence of the direct band gap of InSb due to thermal expansion. The electron-phonon contribution has been ignored. Compare Zollner (SSC, 1991).

Of course, this model for the temperature dependence of the intrinsic carrier concentration is not consistent, because the experimental low-temperature band gap, including the Debye-Waller and self-energy corrections due to zero-point phonon oscillations, was used to calculate the low-temperature masses, but the Debye-Waller and self-energy contributions were ignored in the temperature dependence of the effective masses. To build a consistent $\vec{k} \cdot \vec{p}$ model, we need to remove the zero-point electron phonon contribution to the low-temperature effective masses. We determine a new $\vec{k} \cdot \vec{p}$ momentum matrix parameter as follows:

Within Kane's 8×8 $\vec{k} \cdot \vec{p}$ model, the effective masses of the electron and light hole bands depend on three parameters, the band gap E_0 , the spin-orbit splitting Δ_0 ,

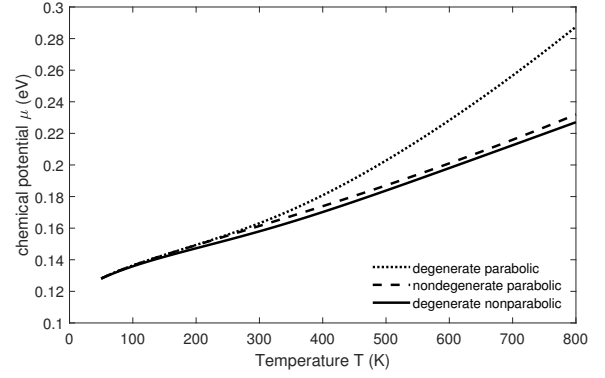


FIG. S24. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap.

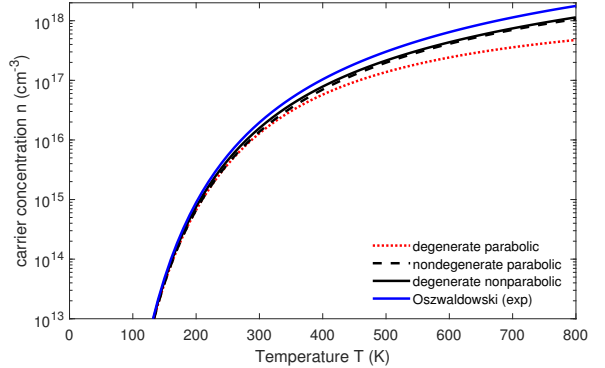


FIG. S25. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

and the dipole matrix element parameter E_P through Eqs. (S31) and (S66). These expressions are accurate within the 8×8 model. Instead of the experimental band gap, we must use the unrenormalized band gap $E_0^u(0K) = 0.243$ eV, see Eq. (29), to calculate the effective masses. Starting with the light hole effective mass

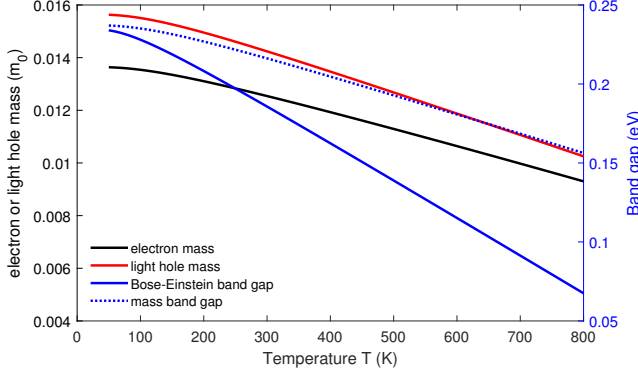


FIG. S26. Effective masses of the electron (black) and light hole (red) bands and band gap (blue) of InSb as a function of temperature, calculated taking into account only the thermal expansion contribution to the band gap. The experimental gap and the thermal expansion contribution to the band gap shift are also shown.

$m_{lh}^* = 0.0156$ and solving Eq. (S66) for E_P , we obtain

$$E_P = \frac{3}{2}E_B \left(1 + \frac{1}{m_{lh}^*}\right) = 23.7 \text{ eV}, \quad (\text{S211})$$

somewhat larger than the usual value of 23.1 eV published by Lawaetz (1971). The corresponding effective electron mass with $\Delta_0 = 0.81$ eV

$$\frac{1}{m_e^*} = 1 + \frac{E_P}{3} \left(\frac{2}{E_B} + \frac{1}{E_B + \Delta_0} \right) \quad (\text{S212})$$

equals 0.0136, which is in excellent agreement with the experimental value.

We therefore repeat the calculation of the temperature dependence of the chemical potential and the intrinsic carrier concentration using this revised value of $E_P = 23.7$ eV. A similar approach was taken by V. G. Orlov and G. S. Sergeev (*Numerical simulation of the transport properties of indium antimonide*, Physics of the Solid State **55**, 2215-2222, 2013), but they do not describe how the use of the unrenormalized gap affects the value of the momentum matrix element. Masut (2022) also uses a similar approach.

The chemical potential and intrinsic carrier concentration of InSb as a function of temperature are shown in Figs. S27 and S28. Compare Fig. 1 in Masut (2022). The agreement between the experimental data of Oszwaldowski and Zimpel (1988) and our model is better than it should be. We need to remember that the large spin-orbit approximation overestimates the electron mass, because Δ_0 is taken to be infinite. Therefore, this model overestimates the intrinsic carrier concentration as well. There is some uncertainty in the temperature dependence of the direct gap, the thermal expansion coefficient, and the density of states heavy hole mass (which is taken to be spherical and parabolic). Therefore, perfect agreement between a model and the experimental data should not be expected.

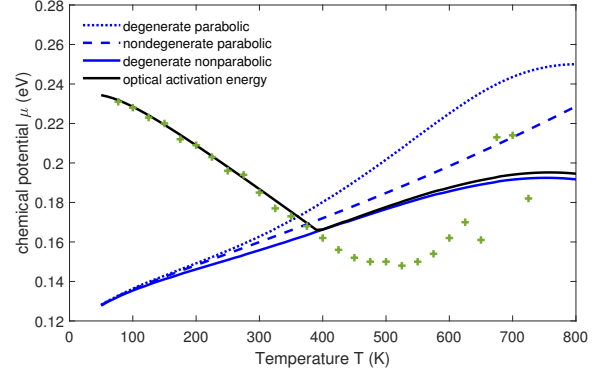


FIG. S27. Chemical potential versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (29) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap. The unrenormalized low-temperature gap was used to calculate the momentum matrix element from the experimental light-hole effective mass. The experimental direct band gap from a fit to the temperature-dependent infrared dielectric function with a Johs-Herzinger parametric oscillator model (symbols) are also shown, compare Rivero Arias (2023).

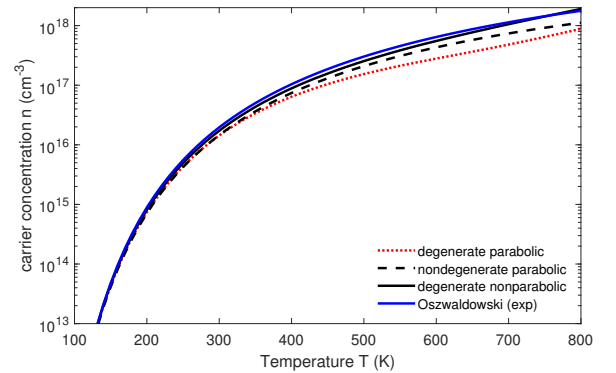


FIG. S28. Intrinsic carrier concentration versus temperature for parabolic bands in the non-degenerate (dashed) and degenerate (dotted) cases. The black solid line shows the degenerate case with the lowest nonparabolic correction in the large spin-orbit approximation. The temperature dependence of the direct gap according to Eq. (S206) was included in the Fermi-Dirac distribution function, but the effective masses were calculated taking into account only the thermal expansion contribution to the band gap. The unrenormalized low-temperature gap was used to calculate the momentum matrix element from the experimental light-hole effective mass. The blue line shows a fit to carrier concentrations determined from Hall measurements by Oszwaldowski and Zimpel (1988).

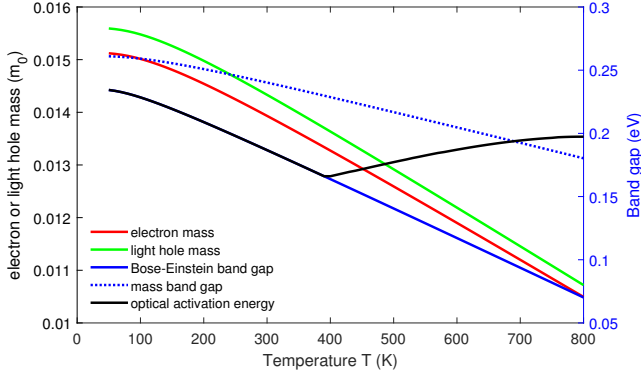


FIG. S29. Effective masses of the electron (black) and light hole (red) bands and band gap (blue) of InSb as a function of temperature, calculated taking into account only the thermal expansion contribution to the band gap. The experimental gap and the thermal expansion contribution to the band gap shift are also shown.

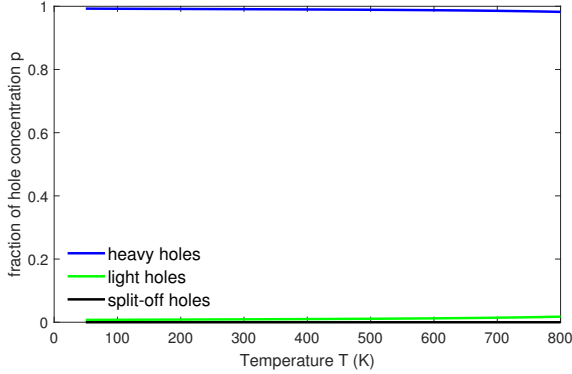


FIG. S30. Fraction of holes in the split-off, light hole, and heavy hole bands, calculated using the expressions in Sec. III B. At high temperatures, the population of the light hole band becomes noticeable on this scale, but still very small.

All of the scenarios described above to calculate the temperature dependence of the effective masses have some flaws, especially how to treat the impact of the zero-point phonon oscillations on the renormalization of the band gap. The most consistent way to handle this matter is described in the main text of this review. The scenarios described in this section can be considered as evidence that treating the band gap as a constant (independent of temperature) or using the full experimental

temperature dependence of the band gap E_0^{exp} to calculate the effective masses is not in agreement with Hall effect measurements.

The optical activation energy (i.e., the band gap observed in an optical absorption or ellipsometry experiment) is increased through the Burstein-Moss shift and given by

$$E_A = \max \left[E_0, E_0 + \left(1 + \frac{m_e}{m_h} \right) (\mu - E_0) \right], \quad (\text{S213})$$

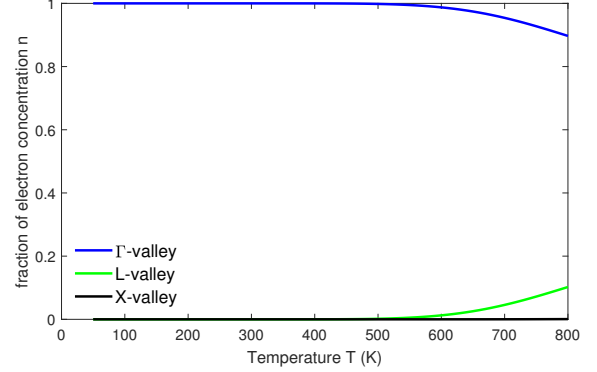


FIG. S31. Fraction of electrons in the Γ -, L -, and X -valleys of the conduction band, calculated using the expressions in Sec. III B. At low and medium temperatures, all electrons are in the Γ -valley, but up to 10% of electrons can be found in the L -valley at the highest temperatures. This was not considered in the analysis of the Hall experiments by Oszwaldowski and Zimpel.⁴⁷

see Chakraborty, Datta, and Ghatak (Physica B **49**, 1179, 2003). The optical activation energy is equal to E_0 if the Fermi level is below the conduction band minimum, but increases as the Fermi level moves into the conduction band. The ratio of the masses takes into account that direct optical interband transitions are not possible at $k=0$, if the Fermi level is larger than the band gap. This optical activation energy needs to be compared to the ellipsometry data given by Rivero Arias *et al.*⁴⁵ as shown in the main text.

S14. SUPPLEMENTARY FIGURES FROM THE MAIN TEXT

The percentages of carriers in the various hole and electron bands as a function of temperature are shown in Figs. S30 and S31.